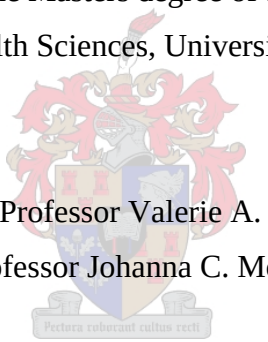
BIOINFORMATICS-BASED STRATEGIES TO IDENTIFY PFHBII-CAUSING AND HCM MAIN LOCUS AND/OR HCM MODIFYING MUTATIONS

Yandiswa Yako

Thesis presented for approval for the Masters degree of Science in Medical Biochemistry at the Faculty of Health Sciences, University of Stellenbosch.



Promoter: Professor Valerie A. Corfield

Co-promoter: Professor Johanna C. Moolman-Smook

December 2004

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and has not previously in its entirety or in part been submitted at any university for a degree.

…….. ……

Signature **Date**

ABSTRACT

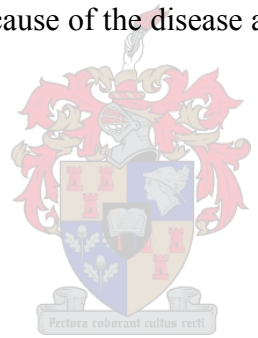
Progressive familial heart block type II (PFHBII) is an inherited cardiac conduction disorder of unknown aetiology, which has been described in a South African family. The disorder was mapped to a 2.9 centimorgan (cM) locus on chromosome 1q32.2-32.3. Clinically, PFHBII manifests cardiac conduction aberrations, that progress to a disease of the heart muscle, dilated cardiomyopathy (DCM). DCM is also reported as an end phase in hypertrophic cardiomyopathy (HCM), another heart muscle disorder. These cardiomyopathies are genetically heterogeneous with some of the genes reported as causes of both disorders. Therefore, genes identified as causes of HCM and DCM were considered plausible candidates for PFHBII mutation analysis. Additionally, this study provided an opportunity to assess potential modifiers of HCM. HCM exhibits marked phenotypic variability, observed within and between families harbouring the same causative mutation.

Genes within the PFHBII locus were selected for PCR-SSCP analysis based on homology to genes previously reported as causing conduction system disorders associated with arrhythmias, DCM and/or HCM. Results were confirmed by direct sequencing and association between the detected variants and HCM parameters was assessed using a quantitative transmission disequilibrium test (QTDT).

Eleven plausible candidate genes were selected within the PFHBII locus and two of the genes, *PFKFB2* and *ATF3*, that encode for 6-phosphofructo-2,6-bisphosphatase (PFK-2/FBPase-2) and activating transcription factor 3 (ATF3), respectively, were analysed for PFHBII-causing and HCM main locus and/or HCM modifying mutations. Mutation analysis of *PFKFB2* and *ATF3* in the PFHBII family revealed no PFHBII causal mutation. *PFKFB2* and *ATF3* were later localised

outside the PFHBII locus, and, therefore, were excluded as PFHBII plausible candidates. Further analysis of the two genes for HCM main locus and/or HCM modifying mutations in the HCM panel identified several sequence variants. QTDT analysis of these variants showed no significant association.

Completion of the Human Genome Project (HGP) and annotation of new genes within the PFHBII locus allowed the identification of more PFHBII plausible candidate genes. Identification of causal mutations in plausible PFHBII candidate genes will allow molecular diagnosis of PFHBII pathophysiology. Furthermore, identification of both HCM-modifying and HCM-causing genes will give insight into the phenotypic variability noted among South African HCM-affected individuals and into the molecular cause of the disease among individuals with HCM-like clinical features.



OPSOMMING

Progressiewe familiële hartblok tipe II (PFHBII) is ’n oorgeërfde hart geleidingsiekte van onbekende etiologie wat in ’n Suid-Afrikaanse familie beskryf is. Die siekte is ’n 2.9 sentimorgan (cM) lokus op chromosoom 1q32.2-32.3 gekarteer. Klinies presenteer PFHBII met geleidingsfwykings wat uitloop op gedilateerde kardiomiopatie (DCM). DCM word ook gerapporteer as ’n endfase in hipertrofiese kardiomiopatie (HCM), ’n ander hartspiersiekte. Die kardiomiopatieë is geneties heterogeen, met ’n aantal gene wat as oorsaak van altwee siektetoestande gerapporteer word. Daarom is alle gene wat geïdentifiseer is as oorsake van DCM en HCM, as moontlike kandidaatgene vir PFHBII mutasieanaliese beskou. Bykomend het hierdie studie die geleentheid gebied om potensiële modifiseerders van HCM te assesseer. HCM toon beduidende fenotipiese variasie binne en tussen families wat dieselfde siekteveroorsakende mutasie het.

Gene binne die PFHBII-lokus is geselekteer vir PCR-SSCP-analiese gebaseer op homologie met gene wat voorheen gerapporteer is om betrokke te wees by geleidingsiesisteemsiektes, geassosieerde arritmieë, DCM en/of HCM. Resultate is bevestig deur volgordebepaling. Assosiasie tusssen ontdekte variante en die siekteparameter is bepaal met ’n kwantitatiewe transmissie disekwilibrium toets (QTDT).

Elf moontlike kandidaatgene in die PFHBII-lokus is geselekteer en twee van die gene, *PFKFB2* en *ATF3*, wat kodeer vir 6-fosfofrukto-2,6-bifosfatase (PFK-2/FBPase-2) en aktiveringstranskripsiefaktor 3 (ATF3) respektiewelik, is vir PFHBII-oorsakende en HCM-hooflokus en/of HCM-modifiseerende mutasies ondersoek. Mutasie-analiese van *PFKFB2* en *ATF3* in die PFHBII-familie het nie ’n siekteveroorsakende mutasie onthul/uitgelig nie. *PFKFB2*

en *ATF3* is later buite die PFHBII-lokus geplaas en dus ook as moontlike PFHBII-kandidate uitgesluit. Verdere ondersoek van díe twee gene vir HCM-hooflokus en/of HCM-modifiserende mutasies in die HCM-paneel het ’n aantal volgorde variante geïdentifiseer. QTDT-analiese van die variante het geen beduidende assosiasies aangetoon nie.

Voltooiing van die Menslike Genoom Projek (HGP) en annotasie van nuwe gene in die PFHBII-lokus het tot die identifikasie van verdere moontlike PFHBII-kandidaatgene gelei. Identifikase van siekte-veroorsaakende mutasies in die moontlike PFHBII-kandidaatgene sal die molekulêre diagnose van PFHBII toelaat en insig in die patofisiologie van die siekte gee. Verder, identifikasie van beide HCM-veroorsakende of HCM-modifiserende gene kan insig gee in die fenotipiese varieerbaarheid wat onder Suid-Afrikaanse HCM-geaffekteerde individue waargeneem word en ook in die molekulêre oorsake van die siekte in individue met HCM-soortige kliniese kenmerke.

ACKNOWLEDGEMENTS

This study was performed in the US/MRC Centre for Molecular and Cellular Biology in the Department of Medical Biochemistry, University of Stellenbosch. I would like to thank the MRC and Doris Crossley fund for their financial support.

I would like to thank God for the strength during the course of this study

My special thanks to my supervisor and mentor, Prof. Valerie Corfield, for her encouragement, enthusiasm, guidance, patience and linguistic support during the course of the study and the final preparation of the thesis. I would also like to thank Dr. Hanlie Moolman-Smook for scientific input, guidance and moral support, you've been quite an inspiration.

I would also like to extend my appreciation to the following people for their input and contribution

- Pedro for mentoring and sharing his expertise.
- Ms. Paula Hedley, Ms. Glenda Dürrheim, Ms. Gift Chauke (my colleague and friend), Ms. Sîan Hemmings, Mrs. Lundi Korkie, Mr. Craig Kinnear, Mr Toy de Lange and Mr. Lester Meissenheimer for their moral and technical support throughout the course of the study.

Finally, special thanks to my father, sisters, brothers and friends for your continued support (moral and financial), encouragement and always believing in me.

INDEX

LIST OF ABBREVIATIONS

α–Tm	alpha-tropomyosin
α-MHC	alpha-myosin heavy chain
β-MHC	beta-myosin heavy chain
λPst	λDNA, digested with the *Pst*I restriction enzyme
A-band	anisotropic band
ACE	angiotensin-1 converting enzyme
ACTC	cardiac actin gene
ADP	adenosine diphosphate
AF	atrial fibrillation
AgNO3	silver nitrate
AMPK	5’ AMP-activated protein kinse
ASD	atrial septal defect
ASREA	allele specific restriction enzyme analysis
ATF3	activated transcription factor 3
ATF3	activated transcription factor 3 gene
ATP	adenosine triphosphate
ATPase	adenosine triphosphatase
AV node	atrioventricular node
AVB	atrioventricular block
BLAST	basic alignment search tool
BLASTN	nucleotide basic alignment search tool
BLASTP	protein basic alignment search tool
bpm	beats per minute
cAMP	cyclic adenosine monophosphate
CHB	complete heart block
CHF	congestive heart failure
cM	centimorgan
CRP3	muscle LIM protein gene
DAP	dystrophin-associated protein
DCM	dilated cardiomyopathy

del/Δ	deletion
dGTP	deoxy-guanine triphosphate
DMD	dystrophin gene
DMSO	dimethylsulphoxide
DOE	Department of Energy
ECG	electrocardiogram
EDTA	disodium ethylenediaminetetraacetate dehydrate
FHC	familial hypertrophic cardiomyopathy
Fn3	fibronectin
G0S2	putative lymphocyte G0/G1 switch gene
HCM	hypertrophic cardiomyopathy
HGP	Human Genome Project
HMM	heavy meromyosin
ICCD	isolated cardiac conduction disorder
I-band	isotropic band
Ig	immunoglobin
IGF-2	insulin-like growth factor-2
IL6	interleukin 6
Ins	insertion
IRF6	interferon-receptor factor-6 gene
IVS	interventricular septum
LGMD	limb-girdle muscular dystrophy
LMM	light meromyosin
LQT3	long QT syndrome 3
LVM	left ventricular mass
LVED	left ventricular end-diastol
LVPW	left ventricular posterior wall
M	marker
MHC	myosin heavy chain
MLP	muscle LIM protein
mRNA	messenger ribonucleic acid
MyBP	myosin binding protein

MyBP-C	myosin binding protein-C
MYBPC3	myosin binding protein-C gene
MyBP-H	myosin binding protein-H
MyBP-X	myosin binding protein-X
MYH7	β-myosin heavy chain gene
MYH6	α-myosin heavy chain gene
MYL2	myosin ventricular regulatory light chain 2 gene
MYL3	myosin ventricular essential light chain 1 gene
NAR	Nucleic Acid Research
NFAT	nuclear factor activated T cell
NIH	National Institute of Health
NIMA-NEK2	never in mitosis gene a-related kinase 2
NKX2.5	homoebox protein Nkx2.5 gene
OD	optical density
P_i	pyrophosphate
PCCD	progressive cardiac conduction disorder
PCR	polymerase chain reaction
PCr	phosphocreatine
PCR-SSCP	polymerase chain reaction-based single strand conformation polymorphism
PFHBI	progressive familial heart block-type I
PFHBII	progressive familial heart block-type II
PFK-2/FBPase-2	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase
PFKFB2	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase gene
PLB	phospholamban
PLN	phospholamban gene
PRKAG2	5’ AMP-activated protein kinase (γ_2 subunit) gene

QTDT	quantitative transmission disequilibrium test
S1	myosin subfragment 1
S2	myosin subfragment 2
SA node	sinoatrial node
SCN5A	sodium channel gene
SDS	sodium dodecyl sulphate
SERCA2a	SR Ca^{2+}-ATPase
SLC30A1/ZNT1	solute carrier family (zinc transporter), member 1 gene
SR	sarcoplasmic reticulum
T	trait
TBE	Tris-borate-EDTA buffer
TGFβ1	transforming growth factor β1
Tn	troponin
TnC	troponin C
TNF-α	tumour necrosis factor-α
TnI	troponin I
TNNI3	troponin I gene
TnT	troponin T
TNNT2	troponin T gene
TPM1	α-tropomyosin gene
TTN	titin gene
TOF	tetralogy of Fallot
VSD	ventricular septal defect
WPW	Wolff-Parkinson-White syndrome

AMINO ACID ABBREVIATIONS

Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid (Aspartate)
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid (Glutamate)
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine
Asx	B	Aspartic acid or Asparagine
Glx	Z	Glutamine or Glutamic acid.

LIST OF FIGURES

FIGURE **page**

LIST OF TABLES

TABLE **page**

CHAPTER 1

INTRODUCTION

INDEX

CHAPTER 1

STUDY OUTLINE

The present study has two main facets, *the first* was the search for progressive familial heart block-type II (PFHBII)-plausible candidate genes within a locus encompassed by markers *D1S70* and *D1S505*, on chromosome 1q32.2-q32.3, to which PFHBII was mapped by P. Fernandez in a previous study (personal communication). PFHBII is a cardiac conduction system disorder that can progress to the cardiac muscle disorder, dilated cardiomyopathy (DCM). Although the genetic cause of PFHBII has not been identified, DCM, which is one of the PFHBII clinical features, is a cardiac muscle disorder that has been widely studied, with several genes implicated as causing the disease.

DCM is not only genetically heterogeneous, but clinical variability has also been noted with some affected individuals also presenting with cardiac conduction system defects (Fatkin *et al*., 1999; Arbustini *et al*., 2002; Itoh-Satoh *et al*., 2002). The conduction system defect has been noted in DCM-affected individuals who harbour lamin A/C (Fatkin *et al*., 1999; Arbustini *et al*., 2002) and titin (Itoh-Satoh *et al*., 2002) mutations. Based on the premise that DCM is one of the clinical features of PFHBII, genes encoding proteins that are structurally and functionally related to previously identified DCM-causative genes may be relevant for PFHBII-causing mutation analysis. Furthermore, DCM-causing mutations have been identified in genes in which certain mutations have been shown to cause hypertrophic cardiomyopathy (HCM). HCM, like DCM, is a cardiac muscle disorder that is also characterised by genetic heterogeneity (Towbin, 2000). Therefore, DCM- and HCM-causative genes and genes that cause cardiac conduction disorders associated with arrhythmias were used as a reference to search for analogous genes that have homologous functional domains and which lie within the PFHBII locus. Bioinformatics tools were used to identify these genes

and other candidate genes that encode proteins which are structurally and/or functionally important in cardiac tissue, which lie within the PFHBII locus. Prioritised candidate genes were then screened for the PFHBII-causative mutation in a panel of individuals from the PFHBII family.

The *second facet* of the study was to investigate if the genes chosen as plausible PFHBII candidates played a role in the development of HCM, or were involved in modifying the phenotypic expression of the disease. Consequently, the selected genes were also screened for HCM main locus mutations and/or HCM modifying variants in a panel of unrelated affected individuals. The HCM panel consisted of HCM-affected individuals with known casual mutations and individuals with HCM-like phenotype. Therefore, screening new plausible candidate genes might elucidate the genetic causes of HCM in those individuals with HCM-like phenotype and the phenotypic variability among affected individuals with identical causal mutations. Plausible PFHBII candidate genes were screened for HCM main locus mutations and/or HCM modifying variants based on the fact that DCM-causing mutations have been identified in genes in which certain mutations were shown to cause HCM (e.g. DCM-causing mutations were also identified in the β-myosin heavy chain-, cardiac troponin T-, cardiac actin-, titin- and α-tropomyosin-encoding genes, all of which are HCM-causative genes) (Regitz-Zagrosek *et al*., 2000; Kamisago *et al*., 2000; Olson *et al*., 1998; Siu *et al*., 1999; Olson *et al*., 2001). Furthermore, HCM-affected indiviuals have been reported to develop DCM-like phenotypes at a later stage of the disease, although it is still unclear whether this in an end-stage HCM phenotype or a transition from HCM to DCM (Seidman and Seidman, 2001).

The first section of the literature review describes the structure of the heart, focusing mainly on the cardiac conduction system and the working cardiac muscle (myocardium). Structural

components of the myocardium, the sarcomere and the cytoskeleton, and their protein complexes will be discussed as abnormalities in these structures have been implicated as causing either DCM or HCM (reviewed by Seidman and Seidman, 2000). Also included in the introduction is a background to the Human Genome Project (HGP), which has provided data and tools that have facilitated identifying genes implicated as causes of genetically inherited diseases (Baxevanis, 2001; Galperin, 2004). The data generated as a result of sequencing and annotating the human genome is integrated and stored in various databases. This information can be assessed and "mined" through the use of bioinformatics, which is also discussed in section 1.2.2.

1.1 The Heart

This section focuses on the cardiac conduction system and the working cardiac muscle (myocardium), which are important components of the heart (Opie, 1991). Each of these components will be discussed, focusing on their structures and how they function.

1.1.1 Structure of the cardiac conduction system

The cardiac conduction system coordinates the production and pathway of the action potential that causes heart beats (Ganong, 1991). It consists of specialised cells that control the rate and rhythm of the heart beat by controlling the rate and direction of electrical impulses as they are propagated through the heart. The adult cardiac conduction system consists of separate morphological components (Fig. 1.1): the sinoatrial (SA) node and the atrioventricular (AV) node, which perform distinct functions (Davies *et al.,* 1983; Ganong, 1991; Moorman *et al*., 1998). The SA node is located in the upper part of the right atrium and is considered to be the pacemaker because of the role it plays in initiating conduction of impulses. The AV node is located lower in the right atrium and forms the slow-conducting myocardial region. Other components include the AV bundle (bundle of His), the right and left bundle branches, and

the peripheral Purkinje fibre network, which form the fast-conducting pathway (Davies *et al*., 1983; Ganong, 1991; Moorman *et al*., 1998).

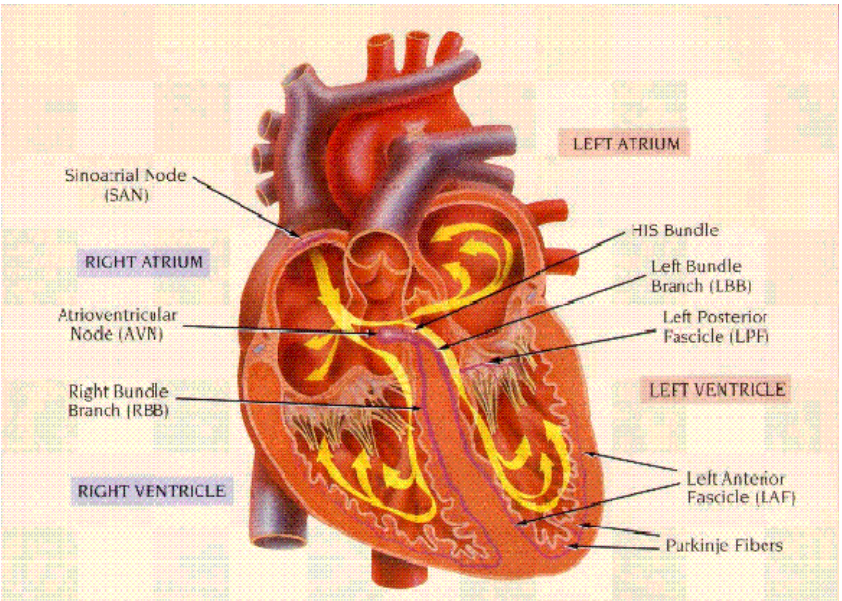


Figure 1.1 Schematic representation of a heart. *Components of the cardiac conduction system and conduction of impulses (shown in yellow) are indicated. The conduction system consists of the following: the sinoatrial node (SA), a branch extending from the sinoatrial node to the upper part of the left atrium, the atrioventricular node (AV), the bundle of His which branches into the left and right bundle branches located each side of the interventricular septum, and small branches called Purkinje fibres. Taken from http://www.cotc.edu/vstone/webdocek.html.*

1.1.2 Propagation of impulses

The SA node can either receive signals from the rest of the body or automatically self-activate to initiate impulse conduction (Wahl, 1999). It receives signals from the body via hormones from the blood and via neurons from the sympathetic and parasympathetic nervous system, acting in a modulatory function only. Once the impulses are initiated in the SA node, they spread quickly through the right atrium via muscle cells in the myocardium (Fig. 1.1). The impulses are transmitted through the gap junctions in the intercalated disks where the cytoplasm of one cardiac cell (cardiomyocyte) connects to, and is continuous with, the cytoplasm of an adjacent cardiomyocyte. Impulses reach the AV node and are delayed for

about one tenth of a second, long enough to allow the atrium to contract and add blood to the ventricle, before the ventricle gets the signal to contract (Ganong, 1991; Rhoades and Pflanzer, 1996).

Impulses travel from the AV node through the bundle of His towards the ventricles. The bundle of His divides into thinner branches, the left and right bundle branches and the bundle branches divide further into thinner branches called the Purkinje fibres, which transmit impulses to all parts of the ventricles. From the Purkinje fibres, the impulses travel from one myocardial cell to the adjacent cell of the ventricular muscle layer. The cells in the lower part of the ventricle contract slightly earlier than the cells in the upper part and this allows blood to be ejected from the ventricles. The process normally takes a fraction of a second, with all cells contracting nearly simultaneously, and it can be monitored by an electrocardiogram (ECG) (Fig. 1.2), which measures the path of impulses as they travel through the CCS and is performed by applying electrodes to the skin.

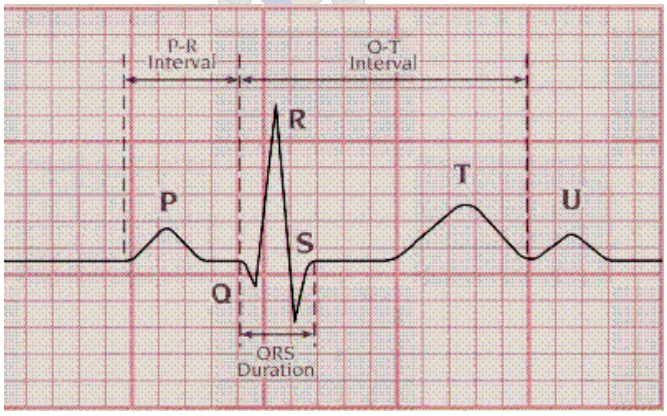


Figure 1.2 A diagram of a typical electrographic recording. *The action potential on different regions of the cardiac conduction system is represented by different 'waves'. Taken from http://www.cotc.edu/vstone/webdocek.html.*

1.1.3 Structure of the working myocardium

This section is focused on the structure of the muscle fibre, highlighting the proteins comprising the contractile unit (sarcomere) and its surrounding proteins, which make up the cytoskeleton.

The working myocardium consists of muscle fibres (myofibrils) that branch and interdigitate, each fibre forming a complete unit surrounded by a cell membrane called the sarcolemma. The sarcolemma forms a tubular network which extends to form an extensive network called transverse or T-tubules (Fig. 1.3) (Ganong, 1991; Opie, 1991). Cardiac myocytes are connected to each other end-to-end by a highly specialised and altered sarcolemma (intercalated discs) (Ganong, 1991; Rhoades and Pflanzer, 1996; Opie, 1991). Intercalated discs are specialised cell-cell junctions that form regions of low electrical resistance, hence providing mechanical and electrical coupling between cells. They are composed of adherens junctions (containing N-cadherin, catenins and vinculin), desmosomes (containing desmin, desmoplakin and desmoglein) and gap junctions (containing connexins) (Stromer, 1995). The nucleus of each cardiomyocyte is centrally located but some have several nuclei that are scattered between myofibrils and immediately beneath the sarcolemma. Cardiomyocytes also contain many mitochondria which generate the heart's energy in the form of adenosine triphosphate (ATP) (Opie, 1991).

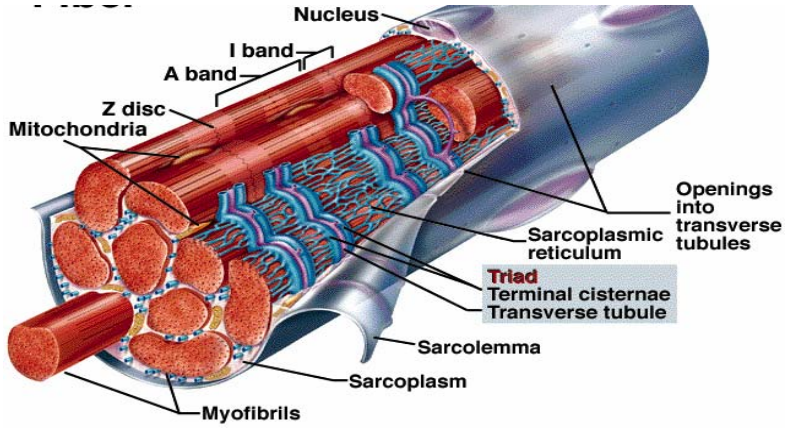


Figure 1.3 Schematic representation of a muscle fibre. *This diagram shows the arrangement of myofibrils and other components of the muscle fibre. Taken from Saladin, 1996.*

The myofibrils, which are the contractile elements of the cardiomyocytes, are cylindrical in shape and 1-2μm in diameter (Fig. 1.3). They are striated in appearance and this is caused by the presence of repeating contractile units (sarcomeres) in each myofibril (Ganong, 1991; Opie, 1991). When viewed under the microscope, each sarcomere contains a series of broad light and dark bands. The sarcomeres are separated from one another by Z lines (Z bands or Z disks) which appear as a dense line in the centre of each light band (Opie, 1991). The Z-disc proteins include α-actinin, filamin, nebulette, telethonin, and myotilin.

The sarcomere consists of two chief contractile filaments, namely, the thick and thin filaments (Fig. 1.4) (Rhoades and Pflanzer, 1996). The thick filaments consist of myosin, myosin-binding proteins C (MyBP-C), H (MyBP-H), and X (MyBP-X), with myosin being the primary structural component (Huxley and Brown, 1967). The thin filaments consist mainly of actin which is attached to the Z disk at either end of the sarcomere and held in a square lattice arrangement by other proteins including alpha (α)-actinin (Alberts *et al*., 1994). The actin is associated with accessory proteins which include α-tropomyosin (α−Tm), and the

troponin complex, which is composed of troponin C (TnC), troponin I (TnI), and troponin T (TnT) (Fig. 1.5). Alpha-tropomyosin is a rod-shaped molecule that binds in the groove of the actin helix. The troponin complex, on the other hand, is anchored to α-Tm predominantly by TnT, and to a lesser extent, by TnI where TnC interacts with both TnT and TnI (Alberts *et al*., 1994) (Fig 1.5). Thick and thin filaments overlap in some regions of the sarcomere forming an anisotropic (A)-band with isotropic (I)-bands on either side, which contain only actin filaments. In the centre of the A-band, there is a clear zone known as the H zone which contains myosin only. Each H zone contains a central dark region, the M line, which contains titin and other M-line proteins (Trinick and Lowey, 1977) (Fig. 1.3 and Fig. 1.4).

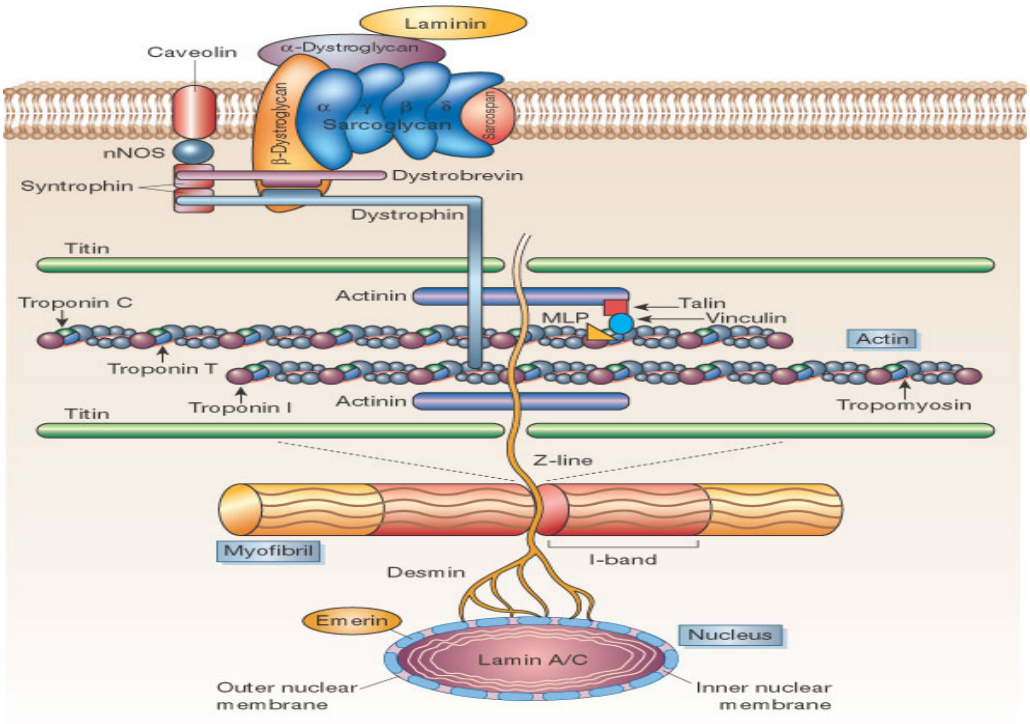


Figure 1.4 Schematic representation of cardiac myocyte structure. *The sarcomere consists of interdigitating thick and thin filaments and is the fundamental structural and functional unit of cardiac muscle. Sarcomeres are connected to the sarcolemma, extracellular matrix, and nucleus by a complex network of cytoskeletal proteins. Taken from Towbin and Bowles 2002*

The myosin extends from the centre of the sarcomere in either direction towards the Z disks and is connected to the Z disks by a giant protein, titin, which extends along the thick filament to the H zone (Fig. 1.4) (Opie, 1991). Titin, together with myomesin and nebulin, form the sarcomeric cytoskeleton that supports the thick and thin filaments. Titin gives elasticity to the cardiac myofibrils and also contributes to the sarcomere assembly and organisation (Fürst *et al*., 1988, Whiting *et al*., 1989). Nebulin and titin form a network of fibres associated with myosin and actin (Alberts *et al*., 1994).

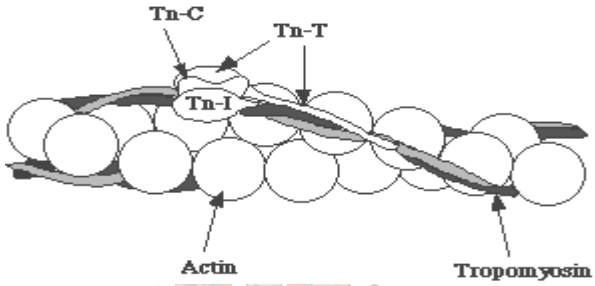


Figure 1.5 Schematic illustration of the troponin complex attachment to α-Tm. *Taken from http://www.embl.de/~wendt/ troponin.html.*

The sarcomeres are connected to each other, the sarcolemma and the neighbouring cardiomyocytes by a complex network of proteins that form the cytoskeleton (Alberts *et al*., 1994; Opie, 1991). The cytoskeleton provides structural support for subcellular structures and transmits mechanical and chemical signals within and between cells. It positions other protein complexes and organelles, and also provides communication paths between them. The cytoskeleton is composed of three types of filaments, namely, microfilaments, microtubules, and intermediate filaments (Alberts *et al*., 1994; Opie, 1991). These filaments constitute different protein subunits: nonsarcomeric gamma (γ)-actin for microfilaments, tubulin for microtubules, and a family of related fibrous proteins such as desmin, vimentin and ankyrin for intermediate filaments. The γ–actin forms a complex network linking the sarcomere via α-

actinin to various components of the costameres. Desmin filaments surround the Z disc connecting adjacent Z discs and forming lateral connections to subsarcolemmal costameres.

The cytoskeletal network in the sarcolemma and the sarcomeric apparatus are linked by costameres, rib-like structures which form subsarcolemmal components (Pardo *et al*., 1983; Koteliansky and Gneushev, 1983; Shear and Bloch, 1985). Costameres are located in a periodic, gridlike pattern, in the level of Z lines and overlying the I bands along the cytoplasmic side of the sarcolemma. They were initially thought to contain vinculin only but were later reported to also include other proteins: talin (Belkin *et al*., 1986), zyxin (Beckerle, 1986), paxillin (Turner *et al.*, 1990), spectrin (Repasky *et al*., 1982; Nelson and Lazarides, 1983; Craig and Pardo, 1983; Messina and Lemanski, 1989), γ-actin (Craig and Pardo, 1983), α-actinin (McGregor *et al*., 1994; Pavalko *et a*l., 1991), fimbrin (Bretscher and Weber, 1980), tensin (Bockholt *et al*., 1992), intermediate filament proteins (Lazarides, 1978; Granger and Lazarides, 1978; Price and Sanger, 1979; Craig and Pardo, 1983), and clathrin (Kaufman *et al*., 1990). These proteins are grouped into three multimolecular complexes: the focal adhesion-type membrane skeleton (vinculin, talin, tensin, paxillin, α-actinin, and zyxin), the spectrin-based membrane cytoskeleton (spectrin, ankyrin, and γ-actin), and the dystrophin/dystrophin-associated glycoprotein (DAG) complex. The DAG complex is the association of dystrophin with a large complex of proteins and glycoproteins called dystrophin-associated proteins (DAP) (Berthier and Blaineau, 1997). The DAP complex consists of one extracellular protein (α-dystroglycan), six integral proteins (β-dystroglycan, α-, β-, γ- and δ-sarcoglycan; and a 25 kDa protein) and three cytoplasmic proteins (α- and β-syntrophins, and dystrobrevin) which are grouped into three subcomplexes, namely, the dystroglycan, sarcoglycan and syntrophin subcomplexes (Yoshida and Ozawa, 1990).

1.1.4 Mechanism of cardiomyocyte contraction

The biochemical basis of muscle contraction involves enzymatic and physical properties of actin, myosin, and six accessory proteins (α-actinin, β-actinin, tropomyosin, troponin complex, cMyBP-C, and M line proteins) of thin and thick filaments (Albert *et al*., 1994). Myosin molecules are long thin fibrous proteins with each molecule consisting of 6 subunits: two myosin heavy chains (MHC) and two pairs of smaller light chains (two essential and two regulatory light chains). The MHC folds into the globular catalytic domain and extends as an α-helical rod through the light chain domain. The globular domain can be proteolytically cleaved into subfragment 1 (S1) which is futher divided into a 25, 50, and 20-kDa fragments, and subfragment 2 (S2). Each myosin head or subfragment 1 (S1) contains a catalytic adenosine triphosphatase (ATPase)- and actin binding sites (Fig. 1.6) (Rayment *et al*., 1993). The S2 subfragment links the two heads and extends into a long rod segment that is bound into the shaft of the thick filament.

Muscle contraction is initiated when the muscle gets signals from the motor nerve which subsequently triggers an action potential in the sarcolemma (Alberts *et al*., 1994). The action potential spreads rapidly into the T-tubules, and is then relayed across a small gap (junctional region) to the sarcoplasmic reticulum (SR). The junctional region contains calcium channels that extend from the SR membrane to make contact with the T-tubule membrane on the other side. In the T-tubule membrane, the voltage-sensitive channels are activated by the action potential, and the calcium channels are triggered to release more calcium ions (Ca^{2+}). Calcium ions are released from their stores in the SR into the cytosol, and myofibril contraction is initiated. During diastole, the interaction between actin and myosin is inhibited by binding of TnI to the actin-Tm complex, but Ca^{2+} release from the SR and its binding to TnC causes a conformational change that weakens the interaction between TnI and actin-αTm, and strengthens TnI-TnC interaction (Solaro and van Eyk, 1996). These changes release the thin

filament from its inhibitory state, inducing actin-myosin interaction and force generation. A reduction in Ca^{2+} concentration ($[Ca^{2+}]_i$) causes a dissociation of Ca^{2+} from TnC and restores the relaxed state.

The myofibril contraction is a result of the interaction between myosin heads and adjacent actin filaments, during which the ATPase in the myosin head hydrolyses ATP (Alberts *et al*., 1994). The ATP hydrolysis, which is followed by dissociation of the tightly bound adenosine diphosphate (ADP) and a pyrophosphate (P_i), produces an ordered series of allosteric changes in the myosin conformation. This results in energy release which is coupled to the movement of myosin head along the actin filament. A "sliding filament" model, was proposed by Huxley in 1957, which explains the generation of force by the sliding movement of the thick filament relative to thin filaments. This movement is mediated by the cyclical attachment and detachment of myosin "cross bridges" to actin. The model was modified to a "swinging cross-bridge" model (Fig. 1.6) (Huxley, 1969; Lymn and Taylor, 1971), which postulates that a swinging movement of the myosin head is a determining factor for its progression along the thin filament. The model that is currently used is the "lever-arm" model, a modification of the "swinging cross-bridge" model (Fig. 1.6) (Rayment *et al*., 1993). The "lever-arm" model states that the catalytic domain remains at a constant position as the elongated light chain domain swings, magnifying conformational changes within the myosin head which accompany ATP hydrolysis. The swinging movement of the myosin tail produces a power stroke. The intermolecular changes of the myosin head (the opening and closing state) link the actin and ATP-binding sites and help the nucleotide pocket to release P_i which signals the onset of the power stroke (Yount *et al*., 1995).

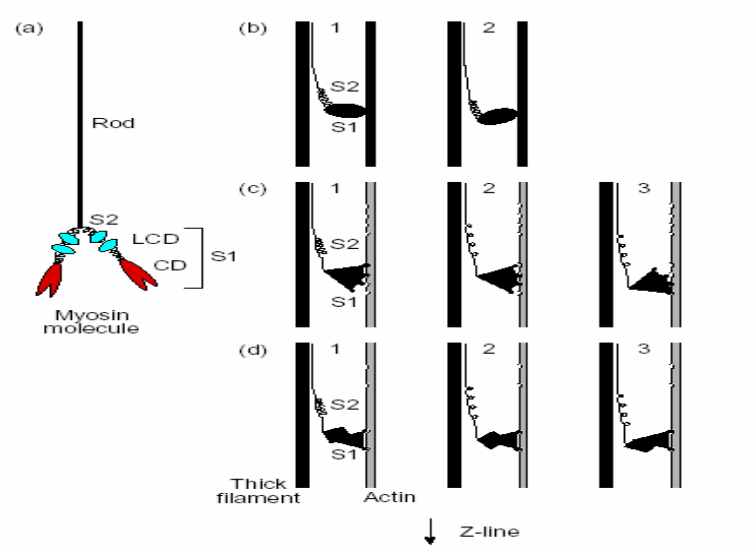


Figure 1.6 Schematic representation of a myosin molecule (a) and alternative early models of swinging crossbridge cycle (b-d) of muscle contraction*. S1 and S2-subfragments 1 and 2, LCD- light chain domain, CD- catalytic domain (see the text for details).* ***b)*** *the "swinging crossbridge" model proposed by Huxley in 1969, (1) a myosin head attaches to actin at a 90° angle, (2) and the whole head swings to a 45° angle.* ***c)*** *A myosin head attaches to actin at one angle (1) and produces force by sliding over the actin binding site (2) to the end of the power stroke (3) (Huxley, 1974).* ***d)*** *The "lever arm model": a myosin head is attached to actin at a stable position (1) and bends in the middle (2) producing force and changing the crossbridge angle (3) (Rayment et al., 1993). Taken from Reedy, 2000.*

Calcium regulates muscle contraction by binding to TnC after it is released from the SR causing a conformational change that weakens the interaction between TnI and actin-tropomyosin, and strengthens TnI-TnC interaction (Solaro and van Eyk, 1996; Bers, 2000). The action potential from the motor nerves to the sarcolemma stimulates Ca^{2+} influx by direct entry into the cell via voltage-gated L-type channels, and, to a lesser extent, via Na^{+}/Ca^{2+} exchanger. Inside the cell, Ca^{2+} binds to cytosolic buffers which include TnC. Calcium influx triggers Ca^{2+} release from the SR to the cytosol through ryanodine and inositol 1,4,5-trisphosphate receptors and, subsequently, myofibril contraction. The extent of cardiac contraction is determined not only by the entry of Ca^{2+} from the extracellular fluid and the SR, but also by the buffering capacity of the cell (Bers, 2000). During relaxation, Ca^{2+} is removed from the cytosol by the SR Ca^{2+}-ATPase (SERCA2a), sarcolemmal Na^{+}/Ca^{2+}

exchanger, sarcolemmal Ca^{2+}-ATPase, and mitochondrial Ca^{2+} unipoter. The $[Ca^{2+}]_i$ is maintained by a complex network of positive and negative feedback mechanisms. The dependence of force on Ca^{2+} is expressed as a function of free $[Ca^{2+}]_i$, but this relationship does not directly indicate the amount of Ca^{2+} required to activate myofilaments, since changes in myofilament sensitivity to Ca^{2+} occur. Changes in myofilament sensitivity are brought about by protein kinase A which phosphorylates TnI. It is also decreased by low pH, and reduced sarcomere length (Bers, 2000).

Calcium also affects the conformation of the thin filament which, as a result, exists in three activation states (Wannenburg *et al*., 2000; Solaro and Van Eyk, 1996; McDonald and Moss, 2000). When Ca^{2+} is not bound to TnC, the thin myofilament is in a "blocked" (off) state and the cross-bridge binding is unable to occur. In a "closed" (off) state, Ca^{2+} is bound to TnC but there is no strong cross-bridge binding; in an "open" (on) state, Ca^{2+} is bound to TnC and cross bridges are strongly bound. According to this model, Ca^{2+} may determine both whether or not cross-bridges bind and the strength of binding. It has also been proposed that changes in Ca^{2+} activation may alter the number of cross-bridges attached, and the kinetics of cross-bridge attachment and detachment (Wannenburg *et al*., 2000; Solaro and Van Eyk, 1996; McDonald and Moss, 2000).

Other regulators of muscle contraction include cMyBP-C and myosin light chains (MLC) (Poetter *et al*., 1996; Sanbe *et al*., 2000; Gautel *et al*., 1995). Cardiac MyBP-C may participate in the adrenergic regulation of cardiac contraction. Ultrastructural studies have shown that phosphorylation of cMyBP-C by cAMP-dependent protein kinase extends myosin cross-bridges from the backbone of the thick filament and changes their orientation. The phosphorylation status of cMyBP-C is a determinant of rigidity and attachment rates of cross-bridges, and Ca^{2+} sensitivity of force production (Kunst *et al*., 2000; Weisberg and Winegrad,

1998). The myosin essential and regulatory light chains (MELC and MRLC) bind to the α-helical lever arm of the myosin cross-bridge and influence force production by modulating cross-bridge kinetics (Poetter *et al*., 1996; Sanbe *et al*., 2000). Numerous extrinsic factors including neurohomoral, endocrine, and haemodynamic factors also influence cardiac contraction *in vivo*.

Calcium dysregulation has been reported to be one of the possible unifying abnormalities caused by HCM-associated mutations in sarcomeric proteins that result in the hypertrophic response. Calcium is predicted to link sarcomere force production to myocyte growth by activating a complex cascade of signalling pathways (Molkentin *et al*., 1998) (section 1.5.4). Thus, any proteins involved in Ca^{2+} regulation and those that are downstream signalling components are plausible candidates for dissection of the molecular causes of HCM, as well as DCM and PFHBII.

1.2 The Human Genome Project (HGP) and Bioinformatics

Sequencing of the human genome has provided insight into various biological processes and serves as a resource of all human genes, and has allowed the identification of genetic causes of many inherited disorders. As this study is focused on identifying genetic causes and modifiers of PFHBII and HCM, respectively, it is relevant to mention the progress in the sequencing of the human genome and integration of its data into various databases.

1.2.1 The HGP

The Human Genome Project (HGP) is an informational collaboration started in the mid 1980s with the primary goal of sequencing the entire human genome (Dib *et al*., 1996). The HGP was officially initiated in the United States (US) by late 1980s under the direction of the US Department of Energy (DOE) and National Institute of Health (NIH). During the course of the

initiative, two groups became competitively involved in the sequencing of the genome using different methodologies and collecting separate data sets (International Human Genome Sequencing Consortium, 2001; Venter *et al*., 2001). The International Human Genome Sequencing Consortium released the first data set from the human genome (draft human genome sequence), which covered most of the genome, for public use in June 2000. The data was made available to the scientic research community through public databases such as the National Centre for Biotechnology Information (NCBI), the Sanger Centre (Ensembl genome browser) and the University of Santa Cruz Centre (UCSC).

The three centres continued to update the human genome sequence data as it was continuously generated (Fig. 1.7). In April 2003, the first finished reference sequence for the human genome was released and consisted of 24 finished chromosomes of 2.9 billion bases and covered about 99% of the gene-containing DNA (Collins *et al*., 2003). Small updates to the assembly were also continuously released after the announcement of the finished reference sequence, as the complex regions were refined and small gaps closed between the large stretches of contiguous sequences (contigs). The finished reference human sequence, along with the results of its analysis and gene annotations, is available for viewing and downloading in various databases, namely, NCBI, the Sanger Centre and UCSC (appendix I).

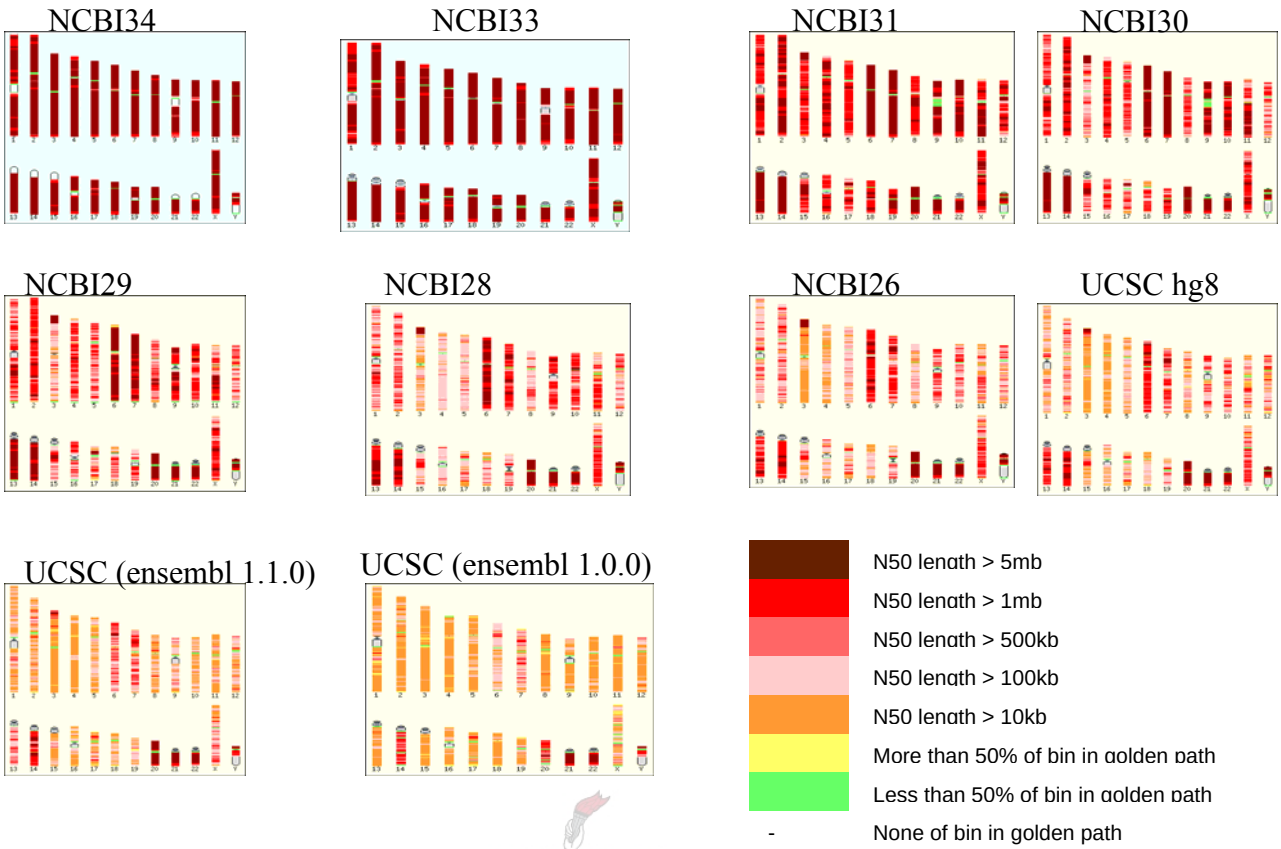


Figure 1.7 Illustration of the progress of the HGP from 12 December 2000 to July 2003*. The regions of the genome are arranged according to the dates they were released: NCBI34- July 2003, NCBI33- 14 April 2003, NCBI31- November 2002, NCBI30- 24 June 2002, NCBI29- 5 April 2002, NCBI28- 24 December 2001, NCBI26- 24 September 2001, UCSC hg8- 6 August 2001, UCSC (ensemble 1.1.0)- 1 April 2001, USCS (ensemble 1.0.0)- 12 December 2000. The completeness of the chromosomal regions is indicated by different colour and these regions are measured by N50 length. Taken from http://www.ensembl.org/Homo_sapiens/stats/status.html.*

1.2.2 Bioinformatics

The data generated through sequencing of the human genome have accumulated over the years (from the time of the initial sequencing), necessitating the development of technologies that can be used to analyse the sequence and apply it in various biological fields. Through the development of these new technologies, which include laboratory techniques and computer programs, biology and computer were combined to form one discipline termed bioinformatics. The new technologies (also referred to as bioinformatics tools) were integrated into various databases according to a particular biological segment that needs to be

addressed. Approximately 548 databases are collectively listed in Nucleic Acid Research (NAR) issues, and the list, together with the short summary and updates of each database, is also freely accessable through the NAR website at *http://www.nar.oupjounarls.org* (Baxevanis, 2001; Galperin, 2004).

Three sequence data centres, NCBI, Ensembl and UCSC, were used in the present study to access published information about the genetically inherited disorders under study, and genes from which the plausible candidates were selected for the analysis of PFHBII-causing and HCM-causing and/or modifying mutations. These three centres are focusing on the detailed annotation of the human genome. In addition, these databases also contain a vast amount of additional annotations for features such as repetitive DNA, expressed-sequence tags (ESTs), CpG islands, and single nucleotide polymorphisms (SNPs). Programs are available within these centres that facilitate data input, quality control, data retrieval, presentation, sequence comparison and functional analysis.

1.3 Cardiac conduction system disorders

1.3.1 Progressive Familial Heart Block-type II (PFHBII)

The first facet of the study involved the search for the gene which causes PFHBII, a disease that was originally described as a cardiac conduction system disorder of unknown aetiology, which is inherited in an autosomal dominant manner. The disorder was identified in a South African Afrikaner family (Fig. 1.8) whose ancestors were traced to a Dutch immigrant who arrived in South Africa in 1713 (Brink and Torrington, 1977).

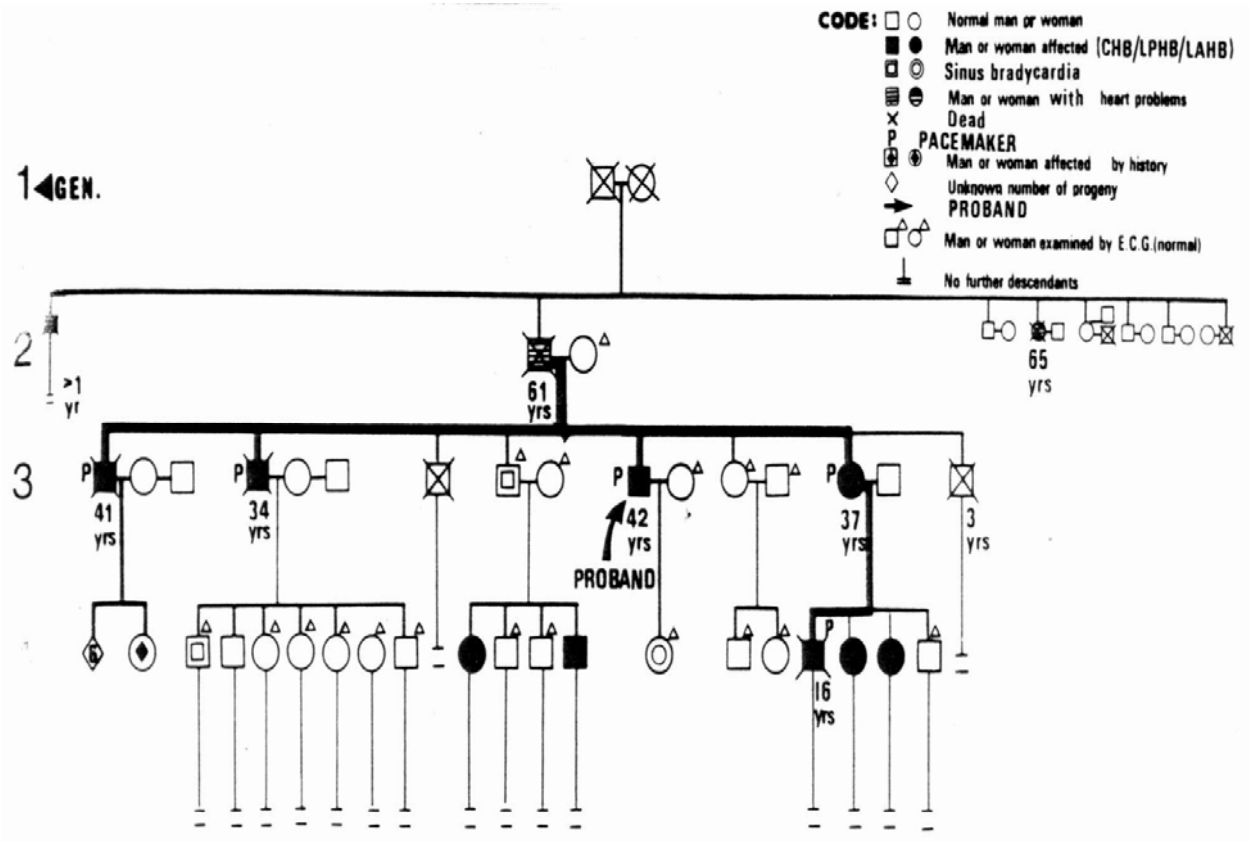


Figure 1.8 A four-generation pedigree in which PFHBII was identified and described (Brink and Torrington, 1977). *Twenty four members of the family were clinically evaluated in the original study and 10 of them were affected.*

The disease was initially characterised by isolated sinus bradycardia (SB), isolated left posterior hemiblock associated with syncopal episodes, Stokes Adams seizures (light headedness) or complete heart block with narrow QRS complexes (Brink and Torrington, 1977). Among members of the PFHBII family that were evaluated in the initial study and additional members who were examined in a later study, several individuals presented a phenotype with a characteristic progression from conduction defects to DCM, demonstrating that cardiomyopathy is one of the clinical features of the disorder (Fernandez *et al*., 2004). The genetic cause of PFHBII has not been identified yet, but linkage studies performed in the affected family mapped the disorder to a locus on chromosome 1q32.2-32.3 that is encompassed by markers *D1S70-D1S505* (Fig. 1.9) (P. Fernandez, personal communication).

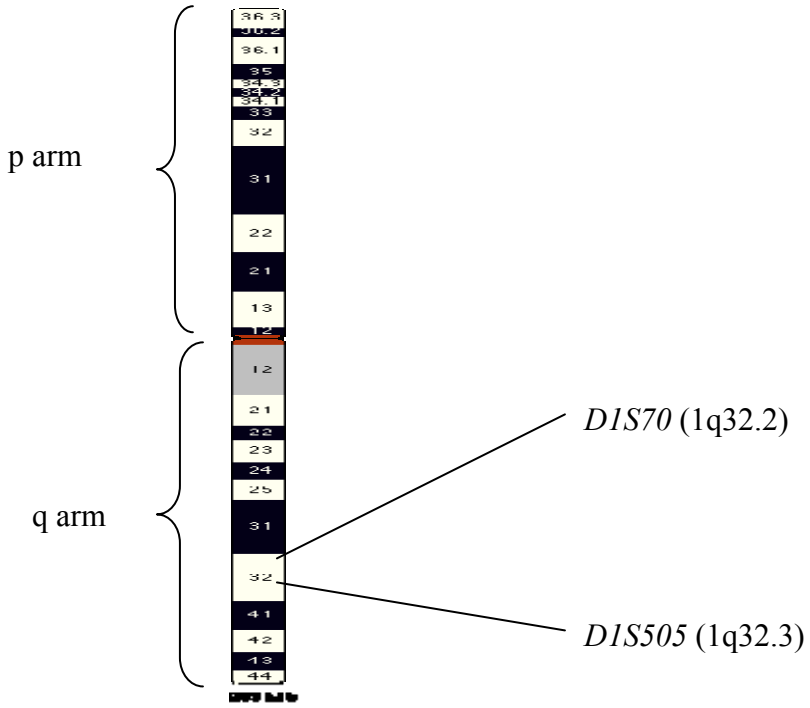


Figure 1.9 An ideogram of chromosome 1. *The PFHBII locus is flanked by markers* D1S70 *and* D1S505. *The markers are localised at position 1q32.2-q32.3.*

Other familial forms of inherited cardiac conduction system disorders have been described, which may give an insight into the pathogenesis of PFHBII, and these include progressive cardiac conduction disorder (PCCD) or Lenègre disease (Lev *et al*., 1970; Lenègre and Moreau 1963), non-progressive cardiac conduction disorder (CCD) (Tan *et al*., 2001), progressive familial heart block type I (PFHBI) (Brink and Torrington, 1977) and isolated cardiac conduction disorder (ICCD) (Stephan, 1978).

PFHBI and ICCD

PFHBI (MIM 113900) was identified in a Caucasian South African family and described as an autosomal dominantly inherited cardiac bundle-branch disorder characterised by RBBB that may progress to complete heart block (Brink and Torrington, 1977; Van der Merwe *et al*., 1986). ICCD is characterised by either a complete or incomplete RBBB with or without QRS frontal-axis deviation and was identified in a Lebanese family (de Meeus *et al*., 1995). In

1995, Brink and co-workers reported linkage analysis, which mapped PFHBII to chromosome 19q13.3. In the same year (1995), de Meeus et al., reported mapping ICCD to chromosome 19q13.3, a region that overlaps with the PFHBI locus. Consequently, the groups (Brink and Stephen, respectively) who described the two disorders formed a brief collaboration and suggested that PFHBI and ICCD are likely to be one clinical entity caused by the same, as yet, unidentified gene.

PCCD and non-progressive CCD

PCCD (MIM 600163) is characterised by progressively deteriorating cardiac conduction through the bundle of His-Purkinje system causing a LBBB or RBBB that eventually leads to complete AV block and sudden death (Lev *et al*., 1970; Lenègre and Moreau 1963). Little was known about the molecular cause of PCCD until Schott and colleagues (1999) performed a linkage analysis in French and Dutch families with PCCD and reported a causative locus on chromosome 3p21. Subsequently, they reported causative mutations identified in a Na^+ channel gene (*SCN5A*), which plays a role in cardiac excitability by allowing Na^+ ion influx that further depolarises excitable myocytes (Hodgkin *et al*., 1952). Interestingly, a few years later Tan et al., (2001) identified a mutation in *SCN5A* that causes the non-progressive CCD that is characterised by broad P waves and QRS complexes, first-degree heart block and severe SB. These findings by Schott et al., (1999) and Tan et al., (2001) were interesting as other mutations in *SCN5A* were reported to cause three types of inherited arrhythmias known as long QT type 3 (LQT3) (Wang *et al*., 1995), Brugada syndromes (Bezzina *et al*., 1999; Naccarelli *et al*., 2002; Baroudi *et al*., 2001) and idiopathic ventricular fibrillation (IVF) (Chen *et al*., 1998).

LQT3, Brugada syndromes and IVF

LQT3 (MIM 603830) is a disorder of repolarisation that is characterised by prolongation of the QT interval on the surface ECG, relative bradycardia, T wave abnormalities, and episodes of ventricular tachyarrhythmias (Wang *et al*., 1995). Brugada syndrome (MIM 601144) is characterised by the elevation of the right precardial ST-segment and ventricular fibrillation (Brugada and Brugada, 1992). Although Brugada syndrome is an arrhythmic disorder, it is accompanied by RBBB, which is indicative of cardiac conduction defects (Schwartz *et al*., 1995; Schwartz *et al*., 1997). Brugada syndrome is also accompanied with ventricular fibrillation, however, a *SCN5A*-missense mutation was identified in a Japanese individual who exhibited IVF but without a typical ECG pattern of the Brugada syndrome (Akai *et al*., 2000). PCCD, LQT3 and Brugada syndrome have been reported to be caused by different mutations in *SCN5A*, however, several studies have demonstrated that some of the mutations can result in either LQT3, the Brugada syndrome or PCCD (Bezzina *et al*., 1999; Veldkamp *et al*., 2000; Grant *et al*., 2002).

Pathophysiology of PCCD, LQT3 and Brugada syndrome

Functional studies of Na^+ channels suggest that Brugada syndrome and PCCD mutations reduce the magnitude of Na^+ current (Dumaine *et al.,* 1999*;* Veldkamp *et al*., 2000; Deschenes *et al*., 2000), whereas LQT3 mutations cause excessive Na^+ current resulting in prolongation of action potential duration and the QT interval (Wang *et al*., 1995; Bennett *et al*, 1995; Priori *et al*., 1999). However, mutations that cause PCCD only result in slow conduction without the ST-T wave changes that characterise Brugada syndrome (Tan *et al*., 2001). Further analysis of the effect of the disease-associated *SCN5A* mutations is required to elucidate the diverse functional behaviour of the channel that may evoke a spectrum of abnormal cardiac excitability, from conduction disease to sudden death.

Mutations in a homeobox protein Nkx2.5 gene (*NKX2.5*), have also been reported to cause an autosominal dominantly inherited congenital heart disease with or without progressive AV block. *NKX2.5* encodes a transcription factor that is important in early stages of heart development, as well as maturation (Chen and Schwartz, 1996; Durocher *et al*., 1997; Tanaka *et al*., 1998). The most common phenotypes that result from *NKX2.5* mutations are AV block and atrial septal defects (ASD), but ventricular septal defect (VSD), tetralogy of Fallot (TOF) and tricuspid valve abnormalities are also included (Benson *et al*., 1999). Several mutations have been reported in *NKX2.5*, which are distributed throughout the molecule (Benson *et al*., 1999; Goldmuntz *et al*., 2001; Ikeda *et al*., 2002). However, the exact mechanisms by which these mutations cause congenital heart disease are still unclear. Some mutations are predicted to reduce the DNA binding affinity of *NKX2.5*, resulting in a loss of transcription activation function of the protein, whereas other mutations decrease the ability of the molecule to form dimers (Kasahara *et al*., 2000; Zhu *et al*., 2000). Further analysis of *NKX2.5* mutations may lead to better understanding of the mechanisms by which the mutations cause congenital heart disease.

1.4 Dilated cardiomyopathy (DCM)

Dilated cardiomyopathy is a cardiac muscle disorder that is characterised by genetic heterogeneity. The disorder is one of the PFHBII clinical features and, as it has been widely studied, DCM gave a platform for investigating a molecular cause of PFHBII. Thus, clinical features and genetic causes of DCM will be discussed in the following subsection.

1.4.1 Clinical characterisation of DCM

Dilated cardiomyopathy is a common complex myocardial disorder characterised by increased ventricular chamber size (dilation) (Fig. 1.10) and diminished systolic function (Dec

and Fuster, 1994; Report by the WHO/ISFC, 1996). Impaired systolic function triggers complex neurohormonal responses that at first are compensatory responses but which ultimately worsen resulting in the onset of heart failure and sudden death, or necessitating cardiac transplantation. Myocyte disarray and increased interstitial fibrosis are not morphological and histopathological features of DCM, as is the case in HCM. Clinical symptoms include shortness of breath, fatigue, intolerance of physical exertion, fainting, light headedness, sweating at rest and sudden death (Abelman and Lorrell, 1989). Although improvement in medical therapy has been reported, DCM continues to be a lethal disease with a 5-year mortality rate of 50 to 75% (Abelman and Lorell, 1989). Epidemiological data indicates that about 36.5 individuals in 100 000 are affected (Gillum, 1986; Michels *et al*., 1992; Grünig, 1998). In approximately 50% of DCM cases the aetiology is unknown, however, myocarditis, toxins, ischaemia, metabolic causes, viral infections, valvular heart disease, coronary heart disease and autoimmunity are recognised as causes. In approximately 25 to 30% of the cases, the disease is caused by inherited gene defects (Grünig *et al*., 1998; Michels *et al*., 1992). However, age-related penetrance and nonpenetrance confound the picture and lead to an underestimation of the prevalence of the condition.

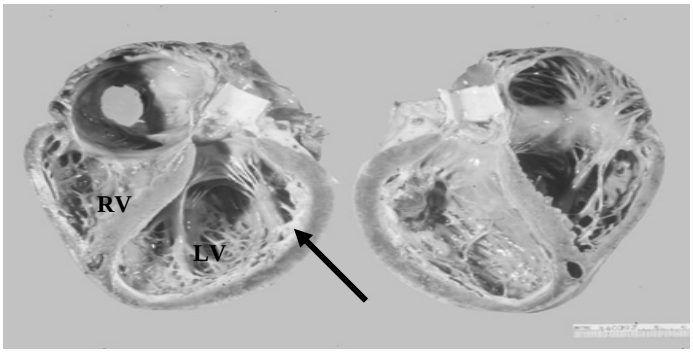


Figure 1.10 A heart from an individual with dilated cardiomyopathy. *The dilated left ventricle is indicated by an arrow. Abbreviations: LV- left ventricle, RV- right ventricle. Taken from www.hopkinsmedicine.org/cardiology/ heart/photographs.html.*

1.4.2 Genetics of DCM

Molecular genetic studies performed in DCM-affected families demonstrated four forms of transmission of the disease: autosomal dominant, autosomal recessive, X-linked and maternal, with the autosomal dominant form being the predominant one affecting individuals of all ages (Richardson *et al*., 1996; Grünig *et al*., 1998; Bowles *et al*., 1996). Linkage analyses provided the first evidence of genetic heterogeneity by identification of several disease loci that were associated with distinct DCM clinical phenotypes such as pure DCM, DCM associated with conduction system disease and/or skeletal myopathy and DCM with sensorineural deafness (Grung *et al*., 1998; Mestroni *et al*., 1999) (table1.1). Subsequently, DCM-causing mutations were identified in thirteen genes that encode cytoskeletal (Li *et al*., 1999; Tsubata *et al*., 2000; Olson *et al*., 2002, Franz *et al*., 2000; Arbustini *et al*., 2000; Muntoni *et al*., 1995; Towbin *et al*., 1993), sarcomeric (Kamisago *et al*., 2000.; Regitz-Zagrosek *et al*., 2000; Olson *et al*., 1998; Olson *et al*., 2001; Siu *et al*., 1999; Gerull *et al*., 2002), mitochondrial (Bione *et al*., 1996; D'Adamo *et al*., 1997; Bleyl *et al*., 1997) and nuclear proteins (Kass *et al*., 1994; Fatkin *et al*., 1999; Bonne *et al*., 1999; Genschel and Schmidt, 2000; Muchir *et al*., 2000).

1.4.3 Mechanisms underlying the pathogenesis of DCM

The mechanisms underlying DCM-causing mutations have not been extensively studied, however, identification of causal mutations in genes encoding cytoskeletal, sarcomeric and nuclear proteins suggested three possible mechanisms: defective force transmission (Towbin *et al*., 1999), defective force generation and disruption of the nuclear structure and function (Seidman and Seidman, 2001; Fatkin and Grahm, 2002).

Table 1.1. *Genetic locus heterogeneity underlying dilated cardiomyopathy and conduction disorders.*

Disorder	Causative gene	Locus	Reference
DCM only			
	TNNT2	1q32	Kamisago *et al*., 2000.
	MYBPC3	11p11	Regitz-Zagrosek *et al*., 2001.
	MYH7	14q12	Kamisago *et al*., 2000.
	ACTC	15q14	Olson *et al*., 1998.
	TPM1	15q22	Olson *et al*., 2001.
	TTN	2q31	Siu *et al*., 1999; Gerull *et al*., 2002
	VCL	10q22-q23	Olson *et al*., 2002
	PLN	6p22.1	Schmitt *et al*., 2003
	DES	2q35	Li *et al*., 1999
	SGCD	5q33	Tsubata *et al*., 2000.
DCM + cardiac conduction system disorder + skeletal myopathy	*LMNA*	1p1-q21	Kass *et al*., 1994; Fatkin *et al*., 1999; Bonne *et al*., 1999; Genschel and Schmidt, 2000; Muchir *et al*., 2000.
X-linked DCM	*DMD*	Xp21	Franz *et al*., 2000; Arbustini *et al*., 2000; Muntoni *et al*., 1993; Towbin *et al*., 1993.
X-linked infantile DCM	*G4.5*	Xq28	Bione *et al*., 1996; D'Adamo *et al*., 1997; Bleyl *et al*., 1997
Brugada syndrome, and LQT3	*SCN5A*	3p21-24	Jiang *et al*, 1994; Wang *et al*., 1995

Abbreviations: *ACTC*- cardiac actin gene; DCM- dilated cardiomyopathy; *DES*- cardiac desmin gene; cardiac dystrophin gene (*DMD*); *G4.5*- cardiac tafazzin gene; LGMD2F- limb girdle muscular dystrophy; *LMNA*- lamin A/C gene; LQT3- chromosome 3-linked Long QT syndrome; *MYBPC3*- cardiac myosin binding protein C gene; *MYH7*- cardiac myosin heavy chain gene; RWS- Romano-Ward syndrome; *SCN5A*- sodium channel, voltage-gated type 5, alpha subunit gene; *SGCD* – δ-sarcoglycan gene; *TNNT2*- cardiac troponin T gene; *TPM1*- cardiac tropomyosin gene; *TTN*- cardiac titin gene; *VCL*- cardiac metavinculin gene.

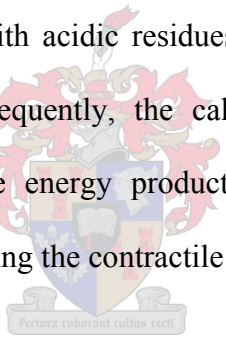
Defective force transmission

The "defective force transmission" hypothesis is based on the fact that DCM-causing mutations were identified in dystrophin, desmin, vinculin and δ-sarcoglycan, which are among structural proteins that form the extrasarcomeric cytoskeleton and the DAG complex (Alberts *et al*., 1994; Opie, 1991). The DAG complex forms part of the subsarcolemmal domains, the costameres, which are the interconnection sites between various cytoskeletal networks that link the sarcomere (contractile unit of the muscle) and the sarcolemma (section 1.1.3). The cytoskeleton transmits force generated during muscle contraction from the sarcomere to the extracellular matrix and protects myocytes from the external mechanical stress (Pardo *et al*., 1983; Koteliansky and Gneushev, 1983). Thus, abnormalities in cytoskeletal proteins and those that form the costameres are proposed to cause DCM through ineffectual propagation of force from the sarcomere to the extracellular matrix and/or resistence to mechanical stress. The pathogenesis may be a result of either a dysfunctional protein or abnormal interaction with other proteins. This concept is supported by the effect of dystrophin mutants on the DAG complex. The truncated dystrophin, when incorporated into the complex, alters the assembly of the sarcoglycan complex (Franz *et al.*, 2000). It has also been noted that two vinculin mutations (Leu954del and Arg975Trp mutations) result in disrupted metavinculin-mediated organisation of actin filaments and abnormal interaction of metavinculin with phospholipids (Olson *et al*., 2002). Furthermore, the location of two DCM-causing mutations on the immobilised end of actin predicts a deficit in actin-cytoskeletal interaction, which may interfere with force transmission from the sarcomere to the extracellular matrix (Olson *et al*., 1998).

Defective force generation

The "defective force generation" hypothesis is based on the identification of DCM-causing mutations in four genes encoding sarcomeric proteins, namely, β-MHC, α-Tm, TnT and cMyBP-C (Olson *et al*., 1998; Kamisago *et al*., 2000; Regitz-Zagrosek *et al*., 2000; Olson *et*

al., 2001; Fatkin and Grahm, 2002). The β-MHC chain is one of the components of the contractile unit, which, by interacting with actin generates force and results in muscle contraction (Huxley and Brown, 1967) (section 1.1.3). Alpha-tropomyosin and TnT are among the proteins that constitute the troponin-tropomyosin complex, a Ca^{2+}-sensitive switch that regulates actin-myosin interaction (Alberts *et al*., 1994). Therefore, any abnormalities in these proteins may interfere with actin-myosin interaction, and, hence, force generation. The concept is supported by identification of DCM-causing mutations in actin-binding domains of both α-Tm and TnT, which, as a result, interefere with actin-myosin interaction (Kamisago *et al*., 2000; Olson *et al*., 2001). Furthermore, atomic structures and biochemical studies have been useful in elucidating the functional effects of TnT mutations (Rayment *et al*., 1995). A deletion of one of the conserved lysine residues was identified in TnT that weakens the complementary ionic interaction with acidic residues in troponin C, the binding partner of TnT (Filatov *et al*., 1999). Consequently, the calcium-stimulated actin-myosin ATPase activity is altered diminishing the energy production for actin-myosin detachment and reattachment, and, as a result, reducing the contractile force within the sarcomere.



Disruption of the nuclear structure and function

Another possible underlying mechanism is disruption of the nuclear structure and function, which is based on the premise that DCM-causing mutations were identified in lamin A/C gene (*LMNA*) (Fatkin *et al*., 1999) that encode components of the nuclear envelop called lamins A and C (Aaronson and Globel, 1975). Nuclear lamins contribute to the structural integrity of the nuclear envelope and provide mechanical support for the nucleus (Gerace, 1986). Thus, DCM-causing mutations identified in *LMNA* are predicted to cause disorganisation of nuclear structure by altering the assembly of dimers and of complex filaments (Fatkin *et al*., 1999; Raharjo *et al*., 2001). Altered transcription regulation is also a possible mechanism that can result in DCM, as predicted by one of the roles of lamins in transcription regulation through

the molecule's binding to chromatin and several transcription factors (Stuurman *et al*., 1998). Further investigation on the mechanisms by which *LMNA* mutations cause human disease will provide insight into the biology of lamins and their importance in the nucleus and viability of cells.

1.5 Hypertrophic cardiomyopathy (HCM)

The second facet of the study involved identification of novel HCM main locus genes and/or HCM modifying genes which may affect the phenotypic expression of this disorder. Phenotypic variability has been noted among HCM-affected individuals carrying identical causal mutations (Fananapazir and Epstein, 1994; Marian *et al*., 1995; Niimura *et al*., 1998). Although diversity in causal mutations has been reported to contribute to this variability, other factors can also influence the disease expression. Therefore, in the following section a summary of HCM-genetic causes, their biophysical properties and clinical outcomes will be given.

1.5.1 Clinical characteristics of HCM

Hypertrophic cardiomyopathy was recognised as early as 1869 (Hallopeau, 1869) but its clinical features were only described in 1950 (Teare, 1958). The disorder is characterised by left and/or right ventricular hypertrophy which is usually asymmetric and can affect different regions of the ventricular wall (Fig. 1.25) (Report of the WHO/ISHF Task force, 1980). The interventricular septum is the most commonly affected part (Wigle *et al*., 1985). Other histological features include myofibrillar disarray and interstitial fibrosis. Myocardial hypertrophy is either a primary disease of the heart, or an adaptive response secondary to haemodynamic disturbances (pressure or volume overload) or endocrine abnormalities.

The prevalence of echocardiographically-defined HCM in a general population in the USA has been reported to be one affected individual in 500 (Maron *et al*., 1995), but in South Africa the prevalence has not been reported. HCM is clinically heterogeneous with inter- and intrafamilial variations ranging from asymptomatic to severe forms with high risk of heart failure or sudden death (Maron, 2000). Genotype-positive individuals may be asymptomatic or present variable symptoms that include shortness of breath (dyspnoea) and chest pains (angina) due to impaired diastolic relaxation of the hypertrophied heart, syncope, embolism, congestive heart failure, deprivation of oxygen (myocardial ischaemia), mitral regurgitation, arrhythmia and outflow tract obstruction (Brachfeld and Gorlin, 1959; Brent *et al*., 1960).

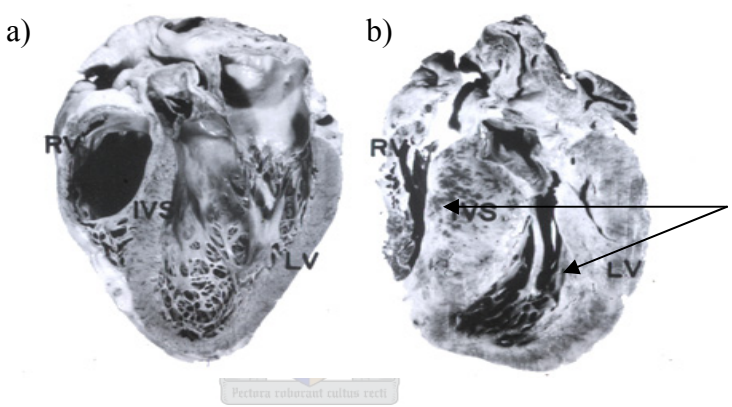


Figure 1.11. ***A longitudinal section of a heart.*** *a) a normal heart and b) a hypertrophied heart. The hypertrophied interventricular septum and left ventricle are indicated by arrows. Abbreviations: RV- right ventricle, IVS- interventricular septum, LV- left ventricle. Taken from www.hopkinsmedicine.org/cardiology/ heart/photographs.html.*

1.5.2. <u>Genetics of HCM</u>

HCM occurs both in sporadic and familial forms, with more than half of the cases being familial (Paré *et al*., 1961). Familial HCM (FHC) is transmitted in an autosomal dominant pattern with one defect in a causative gene reported to be sufficient to cause the disease. Linkage analysis has provided insight into the clinical diversity of FHC with subsequent

identification of disease loci on chromosome 14q12 (*MYH7*) (Jarcho *et al*., 1989), 1q32 (*TNNT2*) (Watkins *et al*., 1993), 15q22 (*TPM1*) (Theirfelder *et al*., 1993), 11q11 (*MYBPC3*) (Carrier *et al*., 1993) (table 1.2). Other genes (*ACTC*, *TNNI3*, *MYL3*, *MYL2* and *TTN*) encoding sarcomeric proteins (cardiac actin, TnI, MELC, MRLC and titin) were also recognised as rare causes of the disease (Mogensen *et al*., 1999; Kimura *et al*., 1997; Poetter *et al*., 1996; Satoh *et al*., 1999). These findings led to a conclusion that HCM is a disease of the sarcomere (reviewed by Seidman and Seidman, 2001). However, HCM-causing mutations were later identified in genes that encode nonsarcomeric proteins, namely, a 5'-AMP-activated protein kinase (AMPK), γ_2–subunit-encoding gene (*PRAGK2*) (Arad *et al*., 2002; Gollob *et al*., 2001), muscle LIM protein-encoding gene (*CRP3*) (cytoskeletal protein) (Geier *et al*., 2001) and phospholamban gene (*PLN*) (sarcoplasmic reticulum protein) (Schmitt *et al*., 2003).

More than 250 HCM-causing mutations have been identified in these genes, which account, in part, for the clinical diversity of the disease. Mutations in the genes listed in table 1.2 account for approximately 50-70 % of all cases of FHC and it is assumed that genes, each accounting for a small proportion, still remain to be identified (Marian and Robert, 1995). Three of the HCM-causative genes predominate, namely β-MHC gene (*MYH7*), TnT gene (*TNNT2*) and cMyBP-C gene (*MYBPC3*), and other genes each account for a minority of HCM cases (Marian and Robert, 1995). Worldwide, 35% of FHC is caused by mutations in *MYH7*, 20% caused by mutations in *MYBPC3*, 15% by mutations in *TNNT2*, and less than 3% by mutations in *TPM1* (Watkins *et al*., 1995). The phenotypic expression of the disease is measured according to the magnitude and pattern of hypertrophy, the age of onset, the incidence of arrhythmias, the histological extent of myofibrillar disarray, the predisposition to heart failure and the incidence of sudden death (Braunwald, 2002). However, according to genetic analysis performed with South African HCM-affected individuals, founder mutations

in *MYH7* (Ala797Thr and Arg403Trp) and *TNNT2* (Arg92Trp) account for 25% 5% and 11%, respectively, of the HCM cases (Moolman-Smook *et al*., 2000).

Table 1.2. HCM-causing chromosomal loci and disease genes

Disorder	Causative gene	Locus	Reference
HCM	*MYH7*	14q12	Geisterfer-Lowrance *et al*., 1990; Vikstrom and Leinwand, 1996.
	MYBPC3	11p11	Bonne *et al*., 1995; Watkins *et al*., 1995.
	TNNT2	1q32	Thierfelder *et al*., 1994.
	TNNI3	19q13	Kimura *et al*., 1997.
	TPM1	15q22	Thierfelder *et al*., 1994.
	MYL2	12q23-q24	Poetter *et al*., 1996.
	MYL3	3p21	Poetter *et al*., 1994.
	ACTC	15q14	Olson *et al*., 2000.
	PLN		
	TTN	2q31	Satoh *et al*., 1999.
	CRP3	11p15	Geier *et al*., 2001.
HCM+WPW+ conduction system disorders	*PRKAG2*	1q36	Blair *et al*., 2001; Gollob *et al*., 2001; Arad *et al*., 2002;

Abbreviations: *ACTC*- *ca*rdiac actin; *CRP3*- cardiac muscle LIM protein gene; HCM- hypertrophic cardiomyopathy; *MYBPC3*- cardiac myosin binding protein C gene; *MYH7*- cardiac myosin heavy chain gene*;* *PRKAG2*- 5'-AMP-activated protein kinase, gamma-2 subunit gene*;* *TNNI3*- cardiac troponin I; *TNNT2*- cardiac troponin T gene*;* TPM1- cardiac tropomyosin gene; *TTN*- cardiac titin gene; WPW- Wolff-Parkinson-White syndrome.

1.5.3 Phenotypic variability of HCM

The phenotypic expression of HCM varies not only with mutations in different genes, but with different mutations in the same gene. This has been noted particularly with mutations in *MYH7*, *TNNT2* and *MYBPC3*. For example, the majority of missense mutations identified in *MYH7* (Arg403Trp, Phe513Cys, Leu908Val and Gly256Glu) are associated with benign clinical and near-normal expectancy, but there are those (Arg403Gln, Arg453Cys and Arg719Trp) that cause a severe form of HCM with early onset that is accompanied by a high risk of outflow obstruction and sudden death (Watkins *et al*., 1992; Anan and Greve, 1994). The benign nature of the Arg403Trp mutation was also noted in a large South African kindred (Posen *et al*., 1995).

Mutations in *MYBPC3* are also associated with a broad spectrum of HCM phenotypes (Erdmann *et al*., 2001). For example, the most severe manifestations of the disease, including life-threatening arrhythmias, have been noted in affected individuals with protein truncations. In contrast, affected individuals with missense mutations generally exhibit a mild form of HCM with late onset of symptoms and a favorable prognosis before the age of 40 years (Charron *et al*., 1998; Niimura *et al*., 1998).

The disease caused by *TNNT2* mutations is usually associated with low penetrance, a relatively mild and sometimes subclinical hypertrophy, but a high incidence of sudden death in the absence of clinical left ventricular hypertrophy (Watkins *et al*., 1995; Moolman *et al*., 1997; Nakajima-Taniguchi *et al*., 1997). However, a different phenotype has been reported in one family with a *TNNT2*-Phe110Glu missense mutation, which presented a more benign form of HCM phenotype with no sudden death, although the hypertrophy itself was overt (Anan *et al*., 1998; Watkins *et al*., 1995). Genetic studies have also revealed that, in some cases, a quarter of genetically affected individuals are clinically asymptomatic, although the

mutant allele is associated with a typical phenotype of the disease in first-degree relatives (Dausse *et al*., 1993; Carrier *et al*., 1997). The degree of functional impairment caused by a particular mutation in a sarcomeric component has been reported to account, in part, for the variability of the phenotypic expression of the mutations (Bonne *et al*., 1998). The functional impairement may vary with the position of the mutation in the molecule, the type of protein involved and modifying factors.

1.5.4 Mechanisms underlying the pathogenesis of HCM

Several mechanisms have been postulated by which HCM-causing mutations in genes encoding sarcomeric proteins may caused HCM. The mechanisms predicted were based on knowledge of the molecular structures of sarcomeric proteins in conjunction with biochemical and biophysical properties of muscle contraction (Seidman and Seidman, 2001). The mechanisms involve either inactivation of the normal allele by the dominant gene mutation, resulting in a reduced amount of the functional protein (haploinsufficiency), or the creation of a mutant protein that interferes with normal function of the protein (dominant negative effect). Most HCM-causing mutations are missense mutations or minor truncations that are unlikely to cause haploinsufficiency. Functional analyses of transcripts with missense mutations demonstrated that the mutant proteins are incorporated into cardiac myofilaments, supporting the "dominant negative effect" hypothesis (Watkins *et al*., 1996; Rust *et al*., 1999). Attempts have been made to explain this concept, but biochemical studies performed demonstrated conflicting results.

Initial biochemical studies of Arg403Gln *MYH7* mutation performed using muscle biopsy specimens from HCM-affected individuals demonstrated a reduced actin filament translocation and force generation when compared with control muscle (Lankford *et al*., 1995; Sata and Ikebe, 1996), whereas, others showed increased actin-activated ATPase activity,

force generation and actin filament sliding velocity (Tyska *et al*., 2000; Lowey, 2002). Furthermore, *in vitro* motility assays performed using embryonic rats carrying the *TNNT2*-In15G>A substitution showed increased velocity of actin-tropomyosin filament translocation and decreased maximum force in the presence of Ca^{2+} (Lin *et al*., 1996; Sweeney *et al*., 1998; Homsher *et al*., 2000). In contrast, functional studies performed with transgenic animal models demonstrated increased Ca^{2+}-sensitivity of force generation caused by *TNNT2*-Arg92Gln and *TNNT2*-Ile79Asn mutations (Morimoto *et al*., 1998). In addition, functional studies performed on HCM-causing mutations in other sarcomeric protein-encoding genes (*MYL3* and *MYL2*) showed increased Ca^{2+} sensitivity (hypercontractility) of force production, thereby increasing force at submaximal Ca^{2+} concentrations (Rayment *et al*., 1995; Fujita *et al*., 1997; Elliot *et al*., 2000; Michele *et al*., 1999). These findings suggest that contractile abnormalities caused by mutations in sarcomeric proteins may not be the unifying dysfunction in HCM, rather, alterations in molecular processes of muscle contractions activate signalling pathways, which trigger myocyte growth and ultimately cardiac hypertrophy (Seidman and Seidman, 2001; Ashrafian *et al*., 2003). Two mechanisms have been proposed that are unifying dysfunctions in HCM, namely, calcium dysregulation and energy depletion (Seidman and Seidman, 2001; Ashrafian *et al*., 2003).

Calcium dysregulation

The release of Ca^{2+} from intracellular stores plays a key role in a wide variety of biological functions such as excitation-contraction, excitation-secretion coupling, as well as transcription and apoptosis (Buratti *et al*., 1995). Calcium regulates muscle contraction by binding to TnC, after it is released from the SR, causing a conformational change that weakens the interaction between TnI and actin-αTm, and strengthens TnI-TnC interaction (Solaro and van Eyk, 1996; Bers, 2000) (section 1.1.4).

The role of calcium dysregulation in HCM was demonstrated in studies of nongenetic causes of HCM, which indicated that angiotensin II mediates growth of cultured myocytes by altering Ca^{2+} concentration (Ito *et al*., 1997). Increased Ca^{2+} has also been noted in rats with surgically induced pressure-overload hypertrophy (Zolk *et al*., 1998). Additional evidence has been noted in functional studies performed with transgenic mice harbouring the Arg403Gln *MYH7* mutation, in which both normal and mutant-containing myocytes were treated with cyclosporin, a calcineurin inhibitor (Fatkin *et al.*, 2000). Calcineurin is a cytoplasmic protein phosphatase that is activated by sustained increase in intracellular Ca^{2+} (Crabtree, 1999). Activated calcineurin dephosphorylates a nuclear factor of activated T cell (NFAT3) transcription factors that causes their translocation to the nucleus where they bind to the zinc finger transcription factor GATA4, leading to activation of embryonic cardiac genes and a hypertrophic response (Crabtree, 2001). In transgenic mouse models, calcineurin inhibition in normal myocytes resulted in increased resting intracellular Ca^{2+} concentration. In contrast, myocytes expressing the myosin mutant fail to elevate intracellular Ca^{2+} in response to calcineurin inhibition, although the development of hypertrophy was accelerated and severe (Fatkin *et al.*, 2000). These data indicate that abnormal Ca^{2+} regulation may play an important role in hypertrophic response in HCM.

Energy depletion

The *PRKAG2* was the first nonsarcomeric protein-encoding gene to be implicated in causing HCM (Arad *et al*., 2001; Gollob *et al*., 2001), suggesting that sarcomeric dysfunction is not the fundamental defect in HCM. The *PRKAG2* mutations that cause HCM raised the possibility that abnormal myocardial energy regulation may be another common mechanism by which HCM mutations result in the development of the disease. The γ_2 subunit of AMPK has been reported to have a sensing mechanism for protecting cells from ATP depletion by regulating diverse intracellular pathways that utilise and generate ATP (Blair *et al*., 2001). It

is activated by cellular stresses that reduce ATP levels and increase AMP, with AMP acting as the main activator interacting with both the autoinhibitory region in the α subunit and the γ–subunit. Upon activation, AMPK switches on catabolic pathways by phosphorylating metabolic enzymes and reducing the activity of non-essential ATP-consuming processes, restoring the energy balance.

The "energy depletion" hypothesis is supported by findings obtained in one study, in which the mean phosphocreatine (PCr)/ATP and P_i/PCr ratios were determined in HCM-affected individuals using P NMR spectroscopy (Jung *et al.*, 1998). The study demonstrated a significantly decreased PCr/ATP ratio and significantly increased P_i/PCr, indicating an abnormal phosphate metabolism. Possible causes of the abnormal phosphate metabolism have been suggested: accelerated work of the heart under resting whole-body conditions, oxygen limitation to cell metabolism severe enough to invoke significant anaerobic contribution to ATP turnover rates and altered carbon and energy sources fueling the cardiac tissue. A transgenic Arg403Gln α-MHC gene (*MYH6*) mouse model also demonstrated similar depletion of PCr/ATP ratios (Spindler *et al*., 1998). Furthermore, increased lipid deposition and mitochondrial pathology was noted in mouse cardiac tissue expressing high levels (92 %) of the *TNNT2*-Arg92Gln mutant, although the sarcomeric structure was not altered (Tardiff *et al.,* 1999). This phenotype strongly suggested that total cellular metabolism (usage of cellular ATP) was altered.

Thus, in their hypothesis, Ashrafian and coworkers (2003) linked Ca^{2+} dyregulation and energy depletion by proposing that energy depletion is the primary stimulus of hypertrophy in HCM, with Ca^{2+} acting as a sensor for cellular energy balance. If there is an imbalance between energy generation and consumption, energy requiring processes such as the SERCA2a pump will be compromised resulting in improper Ca^{2+} re-uptake/removal at the

end of each contractile cycle. Consequently, prolonged cytosolic Ca^{2+} will result to activation of downstream signalling elements such as calcineurin, members of NFAT3 transcription factor, Ca^{2+}/calmodulin-dependent protein kinase and p38 mitogen-activated protein kinase, and ultimately to hypertrophy.

1.5.5 Gene expression profile in both DCM and HCM

Multiple changes in gene expression have been found to occur in response to heart failure, and DCM and HCM are known to be the leading cause of heart failure worldwide (Lim *et al*., 2001; Hwang *et al*., 2002). During the development of heart failure, cardiac myocytes undergo structural and functional changes, that initially, are compensatory but later become deleterious. In an attempt to uncover the mechanisms underlying cellular remodelling as a result of DCM, gene expression profiling was performed in cardiac biopsies of DCM-affected individuals (Yang *et al*., 2000; Grzeskowiak *et al*., 2003). RNA microarray analyses demonstrated overexpression or downregulation of a large number of genes, which is thought to be a secondary response to heart failure. Genes with altered expression profile were categorised into several groups, namely, cytoskeletal and sarcomeric genes, genes involved in cellular energetics, stress protein-encoding genes, genes responsible for protein synthesis and genes responsible for myocardial protein degradation and disassembly (Yang *et al*., 2000; Grzeskowiak *et al*., 2003). Alterations in expression profile of the same group of genes was also noted in cardiac specimens obtained from HCM-affected individuals, although the numbers of transcripts differed in each case. For example, more transcripts that belong to the metabolism subcategory were downregulated in DCM samples than in HCM samples. These studies suggest that the expression level and the number of genes downregulated or upregulated reflect the biological significance of different genes in both HCM and DCM (Hwang *et al*., 2002).

1.6. Genetic modulating factors of the HCM phenotype

One of the characteristics of HCM is the presence of variability in its phenotypic expression, particularly the extent of hypertrophy and sudden cardiac death (Klues *et al*., 1995; Moolman *et al*., 1997; Maron *et al*., 2000). The phenotypic variability is caused, in part, by the genetic heterogeneity and the diversity in causal mutations as highlighted in section 1.5.2. The heterogeneity in the HCM phenotype reflects underlying differences in biophysical properties of mutant proteins, however, it is most likely that other factors, genetic background and/or environmental factors, may also influence the disease expression. This became evident when identical mutations identified in different populations were observed to cause distinct hypertrophic morphologies and variable clinical outcome (Watkins *et al.*, 1992; Fananapazir and Epstein, 1994). For example, the *MYH7*-Arg403Gln mutation that was generally associated with a malignant HCM was reported to have a benign outcome in one Korean family, in which none of the six mutation carriers died at the time of the study (Fananapazir and Epstein, 1994). Similarly, the *MYH7*-Val606Met benign mutation manifested with severe hypertrophy and high incidence of sudden death in four of the eight HCM-affected individuals of one family (Watkins *et al.*, 1992; Fananapazir and Epstein, 1994). Furthermore, phenotypic variability has also been observed among affected individuals of the same family who carry identical mutations. For example, in one family study conducted by Kimura and co-workers (1997), a *TNNI3* Lys183del mutation was identified in a family member who had apical HCM, while another family member with the same mutation had a classical HCM. This mutation has also been reported in other seven families in which only one affected individual presented with apical HCM (Kokado *et al*., 2000).

The influence of genetic background on HCM phenotypic expression has also been investigated and reported in monozygotic and dizygotic twin studies, where a variability in cardiac size was noted irrespective of other factors (Adams *et al*., 1995). This corresponds

with the findings from a family study in which siblings of individuals with LVH were reported to have a higher risk of developing LVH according to the echocardiographic measurements and statistical analyses (Schunkert *et al*., 1999). The contribution of genetic background to HCM phenotypic expression has also been demonstrated in transgenic animal models (Semsarian *et al*., 2001; Sebkhi *et al.*, 1999). In one study performed with two inbred normotensive rat strains, a quantitative trait locus (on chromosome 3) was identified by linkage analysis that affects the difference in cardiac mass of these two strains (Sebkhi *et al.*, 1999). In another study conducted in α-MHC$^{403/+}$ knockout inbred and outbred mouse models, variability in cardiac hypertrophy was observed which was believed to be influenced by genetic factors irrespective of body weight and exercise (Semsarian *et al*., 2001). These studies support the hypothesis that genetic modifiers can influence the response to mutationally altered sarcomere proteins in humans.

Studies performed with different human populations have implicated several genes as candidate modifiers for hypertrophy in HCM, for example, genes encoding components of the renin-angiotensin system, endothelin and its receptors, tumor necrosis factor (TNF-α), chymase, bradykinin and its receptor, aldosterone synthase, insulin-like growth factor-2 (IGF-2), transforming growth factor-β1 (TGFβ1), interleukin 6 (IL6), and platelet activating factor acetylhydrolase (Patel *et al*., 2000; Osterop *et al*., 1998; Deinum *et al*., 2001; Marian *et al*., 1993; Lechin *et al.*, 1995; Yoneya *et al*., 1995; Pfeufer *et al*, 1996; Tesson *et al*., 1997; Yamada *et al.*, 1997). However, analyses of polymorphisms within these genes have shown conflicting results and only explain a small portion of the observed phenotypic variability in HCM. In a study conducted on a Caucasian group of genetically independent HCM-affected individuals, potential modifier genes were investigated for their role in the variability of cardiac hypertrophy (Patel *et al*., 2000). The *TNF-α*-308*G/A* variant was associated with the expression of left ventricular hypertrophy (LVH). The *AA* genotype was associated with

increased LVH when compared with other genotypes. However, functional variants of other candidate modifier genes (*TGFβ1*, *aldosterone synthase gene*, *IL6* and *IGF-2*) that were previously implicated to play a role in cardiac hypertrophy (Yokoyama *et al*., 1997; Hirota *et al*., 1995; Brand and Schneider, 1995; Nezer *et al*., 1999; Jeon *et al*., 1999) were not associated with variable expression of the LVH.

The most commonly studied of these potential modifier genes, *ACE*, that encodes the angiotensin-1 converting enzyme-1 (ACE-1), has given the best supportive evidence of a role in modulating HCM compared to others. An insertion/deletion (*I/D*) polymorphism was identified in *ACE* that was found to be associated with the risk of sudden cardiac death in HCM-affected individuals, with the *DD* genotype more common in HCM families with a high incidence of SCD compared to those with a low incidence of SCD (Marian *et al*., 1993). Subsequent studies have supported this finding by demonstrating an association between the *I/D* genotype and the severity of HCM (Lechin *et al*., 1995; Tesson *et al*., 1997). In one study, the association of the *I/D* genotype was reported to be dependent not only on the particular sarcomeric main locus gene but, also on the causal mutation, and a significant association was observed in HCM-affected individuals with the *MYH7*-Arg403Gln mutation but not in affected individuals with a splice acceptor site mutation in *MYBPC3*. This implies that the trigger for hypertrophy development in response to sarcomeric functional impairment may be influenced by several modifying genes (Tesson *et al*., 1997).

1.7. The present study: Aim and design

The aim of the study was to: i) identify PFHBII candidate genes and to screen prioritised ones for the causative mutation, ii) to screen the PFHBII plausible candidate genes for HCM main locus mutations and/or variants that may modify hypertrophy in HCM. As a strategy to identify PFHBII-causative genes within the PFHBII locus, a catalogue of genes previously reported as

cause of DCM (pure and DCM associated with cardiac conduction system disorder), HCM and conduction system disease was created. Genes in the catalogue were used as a reference to search for genes with homologous functional domains within the PFHBII locus using BLAST similarity search. This is based on the fact that DCM is the primary feature of PFHBII, and DCM-causing mutations were also identified in genes (*MYH7*, *TNNT2*, *ACTC* and *TPM1*) in which other mutations were shown to cause HCM. Furthermore, HCM-affected individuals have been reported to develop a DCM-like phenotype at a later stage, although it is still unclear whether this is an end-stage HCM phenotype or a transition from HCM to DCM (Seidman and Seidman, 2001). This is noted in affected individuals harboring *TPM1*-Glu180Val, *MYH7*-Arg453Cys, *MYH7*-Leu517Met, *MYH7*-Gln734Glu and *MYBPC3*-Gln1012X mutations. This has caused a fundamental paradigm shift with regard to exclusive causes of HCM and DCM (Olson *et al*., 2001; Moolman-Smook *et al*., 2003). Furthermore, HCM and DCM with associated conduction system defects are commonly noted in conjunction with each other. These data may indicate common pathophysiological events leading to conduction disorders, HCM or DCM, and may provide a handle to identify the cause of PFHBII. Therefore, any gene within the PFHBII locus that is structurally and/or functionally related to genes implicated as causing DCM or HCM, is a plausible candidate for PFHBII mutation analysis.

Apart from using BLAST similarity search to identify plausible PFHBII candidates, Ensembl and UCSC gene identification tools were also used. Genes were considered strong candidate genes if they lie within the PFHBII locus, are expressed in the cardiac tissue, and play a structural and/or functional role in the heart. The plausible candidate genes were screened for the possible PFHBII-causing mutation in the PFHBII family using polymerase chain reaction-based single strand conformation polymorphism (PCR-SSCP) mutation analysis and automated sequencing. In addition, plausible PFHBII candidate genes were screened for HCM main locus mutations and/or variants that may either modify the expression of the HCM phenotype or

contribute to the development of the disease. Allele specific restriction enzyme analysis was used to confirm the presence of identified variants and to genotype other individuals in the families of HCM proband. Thereafter, the association between the identified genotypes and the disease phenotype was determined by a quantitative transmission disequilibrium test (QTDT).

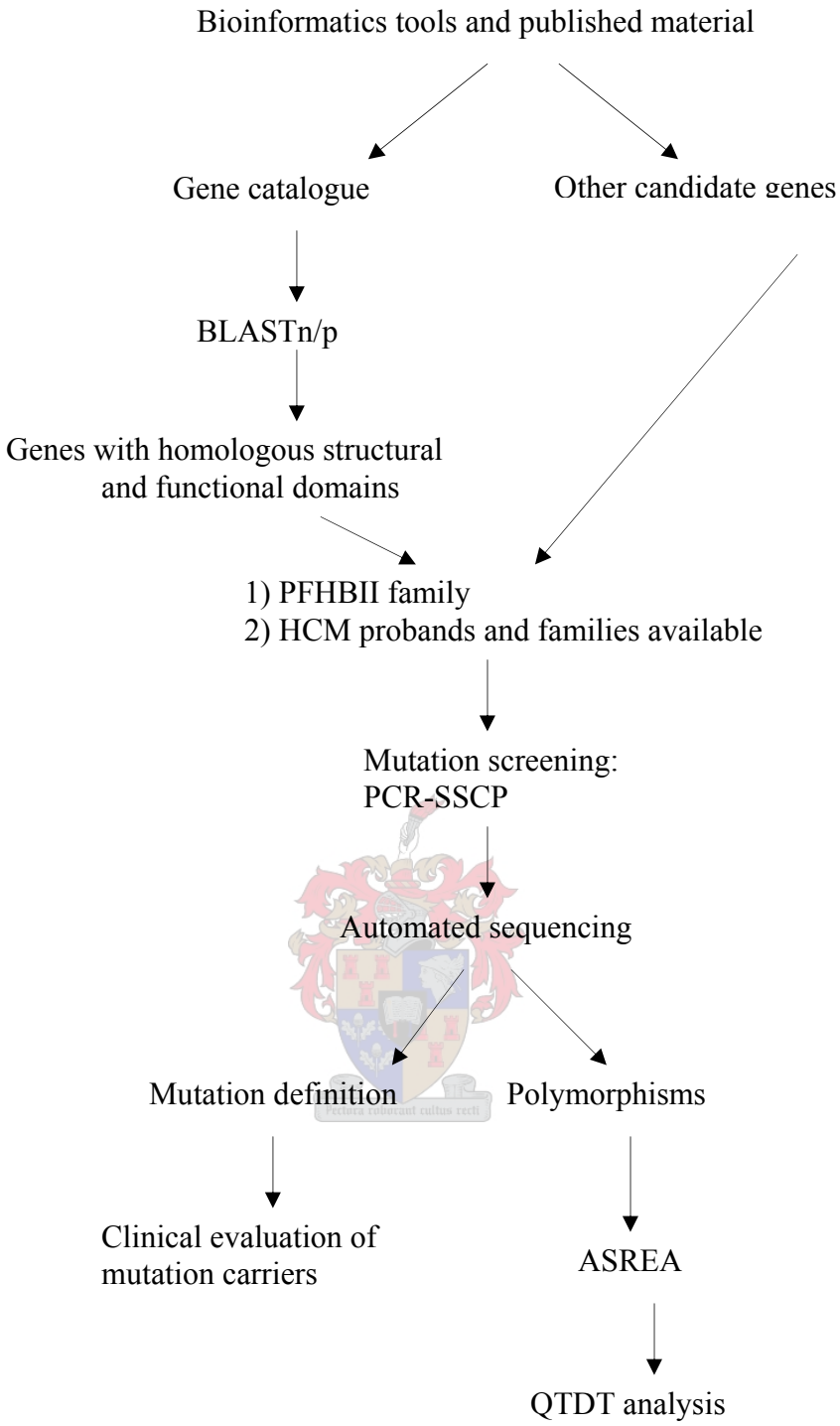


Figure II.13. ***Flow chart summarising the design of the present study***. *Explained in the text. Abbreviations: BLASTn/p- basic alignment search tool for nucleotides and protein sequences, PCR-SSCP- polymerase chain reaction-based single strand conformation polymorphism analysis, ASREA- allele specific restriction enzyme analysis.*

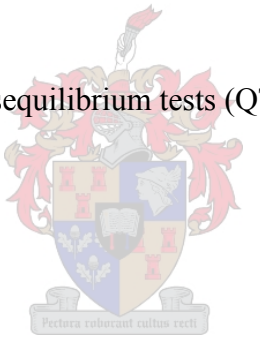
CHAPTER 2

MATERIALS AND METHODS

INDEX

INDEX

page



2.1 Identification of candidate genes within the PFHBII locus

2.1.1 Preparing a catalogue of cardiac conduction system disease-, DCM-, and HCM-causing genes

Database searches (appendix I) and previously published material were used to prepare a catalogue of genes that have been reported to cause CCS disease, DCM (pure or associated with CDDC), or HCM. The chromosomal locus and the accession number of each gene were also included in the catalogue.

2.1.2 Basic Alignment Search Tool (BLAST) search

The discovery of DNA and amino acid sequence homology to a known gene and protein, or protein family, often provides the first clues about the function of a newly sequenced gene. There are a number of software tools for identifying sequence homology between two sequences and BLAST is one of them (appendix I). BLAST programs have been developed to compare protein or DNA queries with protein or DNA databases in any combination, with DNA sequences undergoing conceptual translation before any comparison is performed. The BLASTN and BLASTP option of the NCBI-BLAST programmme compares sequence queries to nucleotide (BLASTN) or protein databases (BLASTP) (Altschul *et al*., 1990).

The two BLAST programs, BLASTN and BLASTP (appendix I), were used in the present study to search for genes within the PFHBII locus with sequences homologous to those of the catalogued genes.

Nucleotide BLAST (BLASTN) search

The BLAST search was initiated in June 2001, at the time, the sequence status of the human genome was in a draft form and gene annotation was incomplete (refer to section 1.2). The draft genome sequence consisted of approximately 25,790 genes (release 0.8.0), including 94% of

known genes, out of 30,000-35,000 genes that were predicted using expressed sequence tags (ESTs), comparative genomics and the completed human chromosomes (Semple *et al*., 2001). Thus, complete *in silico* or electronic versions of coding nucleotide sequences of several genes were available in the NCBI, Ensembl and UCSC databases. Coding sequences of the catalogued genes were obtained (using the "copy" option) from GENBANK (NCBI website) and entered (using the "paste" option) in a FASTA format into the "query" window at the NCBI BLAST or Ensembl BLAST (appendix I) webpage. The "masking" option (also known as the low complexity filter option) was chosen to minimise the influence of the low complexity regions. This option was selected because many regions throughout eukaryotic genomes consist of sequences that code for transmembrane helices or coiled-coil domains, which may, therefore, result in false-positive alignments (Huynen *et al*., 1998). The query sequence was then submitted and the results were accessed in a separate web page. A significant threshold level of 95% or greater, or e-value less that 0.01, was used to select sequences homologous to a "query" sequence of 40 nucleotides or more (Altshul *et al*., 1997; NCBI BLAST tutorial). The algorithm indicates the difference between alignment matches that have only mathematical significance and matches that have a biological significance (Altshul *et al*., 1997)

Protein BLAST (BLASTP) search

BLASTP can be used to find homologous protein coding regions by directly entering the sequence of the protein encoded by the gene to be analysed into the query window. Protein sequences of genes in the catalogue were submitted to BLASTP, as described in the above section. All resultant sequence alignments with an e-value lower than 0.01, or with percentage identity of 40-50%, were considered significant (Altschul *et al.,* 1997).

Initially, the present study searched for functional domains in the PFHBII locus that were homologous to functional domains of proteins in which mutations have been described as causing DCM associated with CCS disorders. CCS disorders complicated with arrhythmia and HCM were also included in the search for the reasons discussed in section 1.7. However, as the study progressed, improvements were made to bioinformatics search tools, which resulted in the use of a modified bioinformatics approach to identify possible PFHBII and HCM candidate genes at the PFHBII locus, as described in the following subsection.

2.1.3 Databases used to identify annotated genes within the PFHBII locus

Newly annotated genes were continuously added to the PFHBII locus as the human genome sequencing progressed. The Ensembl and UCSC databases (appendix I) provide tools for viewing annotated genes within specified genomic regions, their nucleotide and protein sequences, tissue expression profiles and their predicted functions. This information can be accessed in specific databases that are linked to the Ensembl and UCSC websites (appendix I). Therefore, in addition to using the BLAST search of different databases and comparing the results, Ensembl and UCSC gene identification tools were also used.

Ensembl gene identification tools

The Ensembl home page provides links to genomic data of *Homo* (*H*) *sapiens*, *Pan troglodytes*, *Mus musculus*, *Rattus* (*R*) *norvegicus*, *Danio rerio*, *Fugu rubripes*, *Anopheles gambiae*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Caenorhabditis briggsae*, and *Gallus gallus*. The PFHBII locus was viewed in the *contigview* web page by clicking chromosome 1 (see website) and selecting the region flanked by markers *D1S70* and *D1S505*. The *contigview* web page allows the user to scroll along the entire chromosome, and, it allows the user to click on an annotated (*in silico*) transcript (known or computer predicted), which

activates links to other databases that have more detailed information about the gene in question. (Fig 2.1).

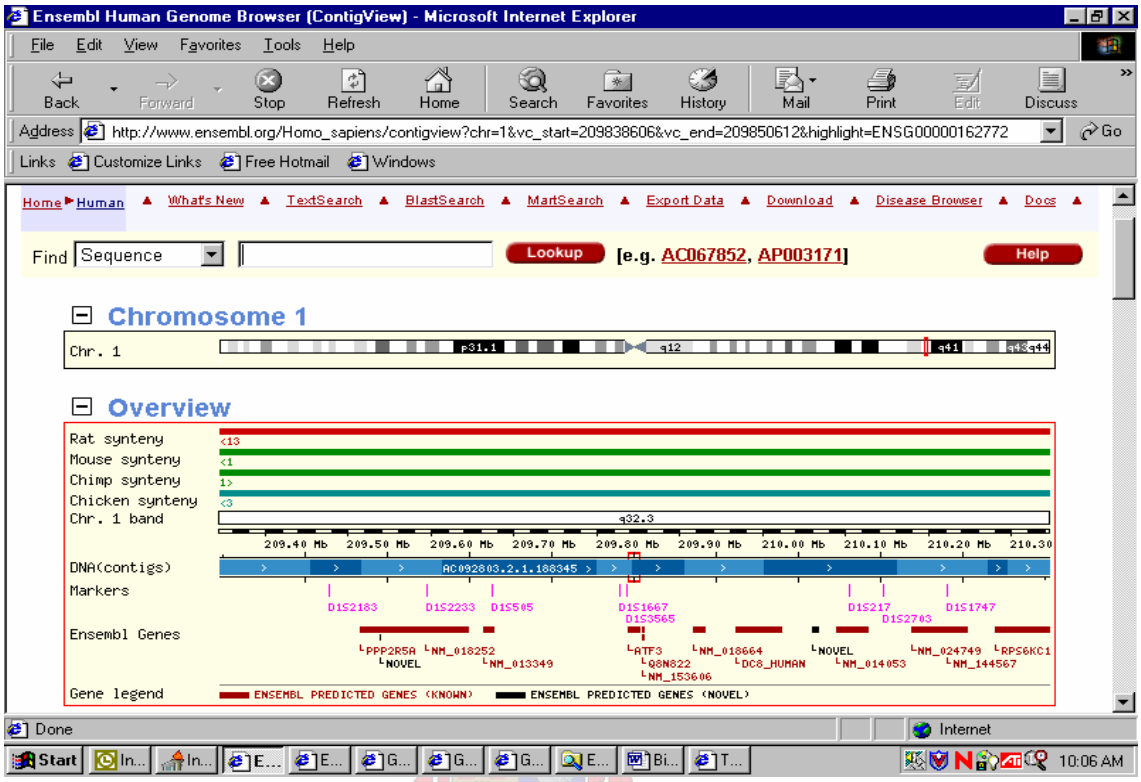


Figure 2.1 A representation of Ensembl contigview showing a region in human chromosome 1. *The region is shown at high resolutions and detailed information on each annotated gene can be accessed by clicking in any of them. Genes are coloured brown. Novel 'Ensembl genes' are shown in black. Taken from Ensembl database (appendix I).*

UCSC gene identification tools

Alternatively, the PFHBII region was accessed via the UCSC browser. In this case, a search for a gene was performed by specifying the chromosomal region, which can be viewed either as a chromosome band or a range of bases (e.g. chr 1: 206860725-209674431bp). GENSCAN gene predictions (known and novel genes), mRNA alignments or interspersed repeats are graphically displayed in the middle of the UCSC web page (Fig. 2.2). Clicking on a transcript accesses information relevant to that particular gene, as all information in the UCSC database is linked to other websites (e.g., GO, GENSCAN). Thus, candidate genes could be selected

using the comprehensive data provided in the UCSC database, as well as information obtained in other specialised databases.

For the present study, it was opted to use the Ensembl and UCSC databases to identify plausible candidate genes at the PFHBII locus. Data obtained from both databases were compared to each other, thereby creating a comprehensive profile of the selected candidate gene. New candidate genes were prioritised according to their tissue expression profile and their possible role in cardiac functioning. Genes selected within the PFHBII locus were screened for PFHBII-causative mutation, as well as for HCM main locus and/or HCM-modifying mutations (as discussed in chapter 1).

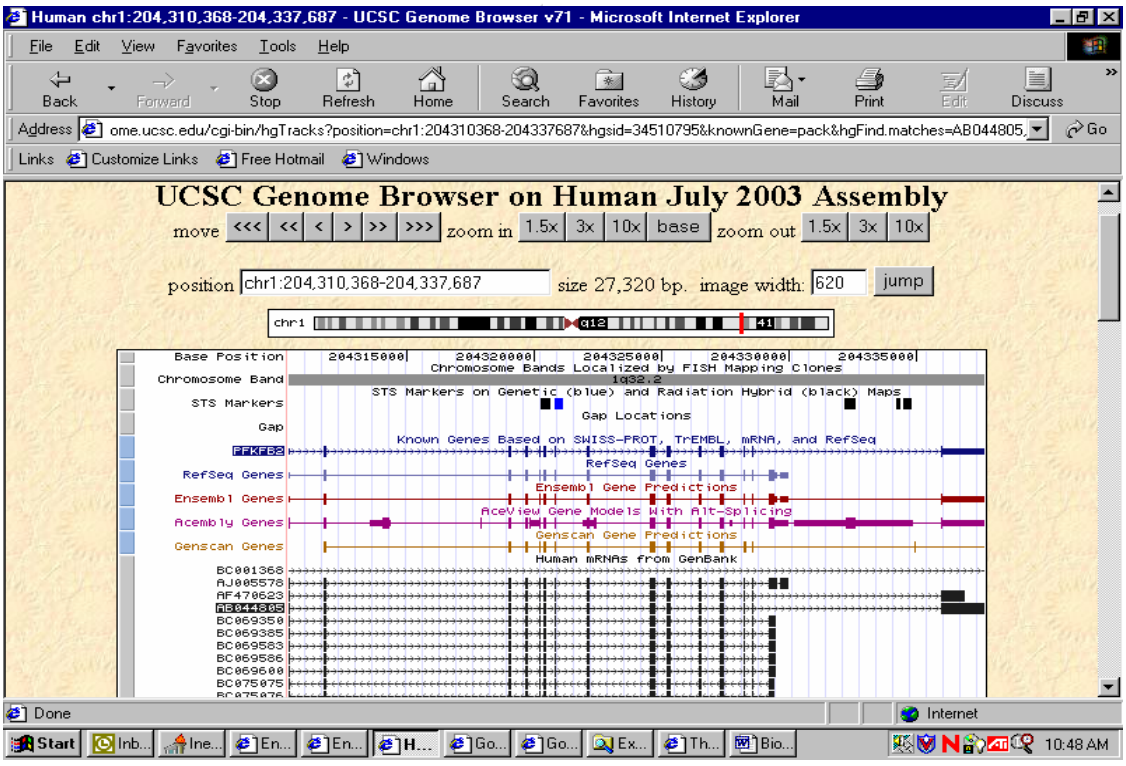


Figure 2.2 A representation of the UCSC main browser display showing a region of human chromosome 1. *The web page contains three main panels: The top panel contains a series of controls that are used for searching, zooming and scrolling across a chromosome. In the middle panel is a graphic display of annotated genes at a locus. Detailed information on the selected gene can be accessed by clicking in any of the genes (see text). Taken from UCSC database (appendix I).*

2.2 Identification of promoter elements of candidate genes

Gene expression profiling has been performed in DCM and HCM-affected individuals and several genes have been reported to have altered expression levels, which may be due to mutations in promoter elements (Yang *et al*., 2000; Grzeskowiak *et al*., 2003) (section 1.4.3.3). For this reason, promoter regions of candidate genes were also screened for possible mutations that may affect the expression level of these genes in affected individuals.

2.2.1 PFKFB2 promoter region

The promoter region of *PFKFB2* is not completely characterised and in order to analyse this region for PFHBII- and HCM-causing and/or modifying mutations, and to predict the effect of mutations in PFHBII and HCM, it was necessary to identify the promoter elements (transcription-binding sites) within this region. Five transcription-binding sites were previously identified within a 500-bp 5' region of *H. sapiens PFKFB2* which corresponds to exon 1 of the gene (Heine-Suner *et al*., 1998). In order to identify the exact positions of these five transcription-binding sites, European Bioinformatics Institute (EBI) alignment tools (appendix I) were used, and positions of three *Sp1* sites and one *HNF1* conserved transcription-binding sites were identified (refer to section 3.2.3).

2.2.2 ATF3 promoter region

The *ATF3* promoter region is well defined and is available in NCBI nucleotide database under the accession number U37542.

2.3 Selection of subjects

Mutation analyses were performed in two groups of subjects, namely, the PFHBII-affected family (Brink and Torrington, 1977; Fernandez *et al*., 2004) in which candidate genes were analysed for PFHBII-causing mutations, and the HCM panel (Moolman-Smook *et al*., 2000), in

which candidate genes were analysed for HCM main locus and/or modifying mutations. The selection of subjects making up the two groups is described below.

2.3.1 PFHBII family

The study was approved by the Ethics committee of the University of Stellenbosch, Tygerberg, and informed vocal and written consent (parental consent in the case of minors) was obtained from all individuals entered into the study. PFHBII has been described in one family only (Fig 2.3) (Brink and Torrington, 1977). Clinical records and family history of 24 members of the PFHBII family were previously collected and filed for a long-term study. These individuals were examined and described in the original PFHBII study (Brink and Torrington, 1977), and have been followed to the present and the clinical data was collected as part of an ongoing study (Fernandez *et al*., 2004). Affected and unaffected individuals that were entered into this author's study are shown in table 2.1.

2.3.2 HCM panel

The study was approved by the Ethics committee of the University of Stellenbosch and informed vocal and written consent (parental consent in the case of minors) was obtained from all individuals entered into the study. The HCM panel was established from South African HCM-affected individuals and individuals with echocardiographic HCM-like features, and forms part of an ongoing study conducted by Dr. Moolman-Smook in the Department of Medical Biochemistry, Faculty of Health Sciences (University of Stellenbosch). The panel consisted of 143 unrelated individuals of different ethnic backgrounds (Coloured, Blacks and Caucasians), 31 with previously identified HCM-causative mutations (*MYH7*-Ala797Thr, *MYH7*-Arg249Gln, *MYH7*-Arg403Trp, *MYH7*-Glu499Lys, *MYH7*-Arg719Trp, *MYH7*-Ala1379Thr, *MyBPC3*-Δc756, *MyBPC3*-Arg654His, *MyBPC3*-Val896Met, *MyBPC3*-*IVS7+1G>A*, *TNNT2*-Arg92Trp, and *TNNTI3*-Arg186Trp) (Fung *et al*., 1999; Moolman *et al*.,

1993; Moolman *et al*., 1995; Moolman-Smook *et al*., 1998; Moolman-Smook *et al*., 1999; Moolman-Smook *et al*., 2000). The panel was screened continuously to identify new causative mutations as the previously identified mutations only account for approximately 22% of the total (Dr. Moolman-Smook, personal communication), and were not found in other affected individuals in the panel.

Clinical information about affected individuals was obtained from Tygerberg Hospital (Cape Town, RSA), clinics, private practises and other hospitals nationwide through collaboration between the University of Stellenbosch and physicians. Family trees, or a preliminary pedigree structure of individuals entered into the study, were traced by Sr Althea Goosen, if the family had a history of the FHC. In addition, a detailed history and clinical records of the probands and their relatives, where possible, were also obtained by Sr Goosen to establish the frequency of sudden deaths and disease-related deaths. Probands with clinical features and a history of HCM, or any individuals who demonstrated HCM-like hypertrophy and symptoms upon examination, were continuously identified, informed of the HCM mutation screening project, and given the opportunity to participate if they so wished.

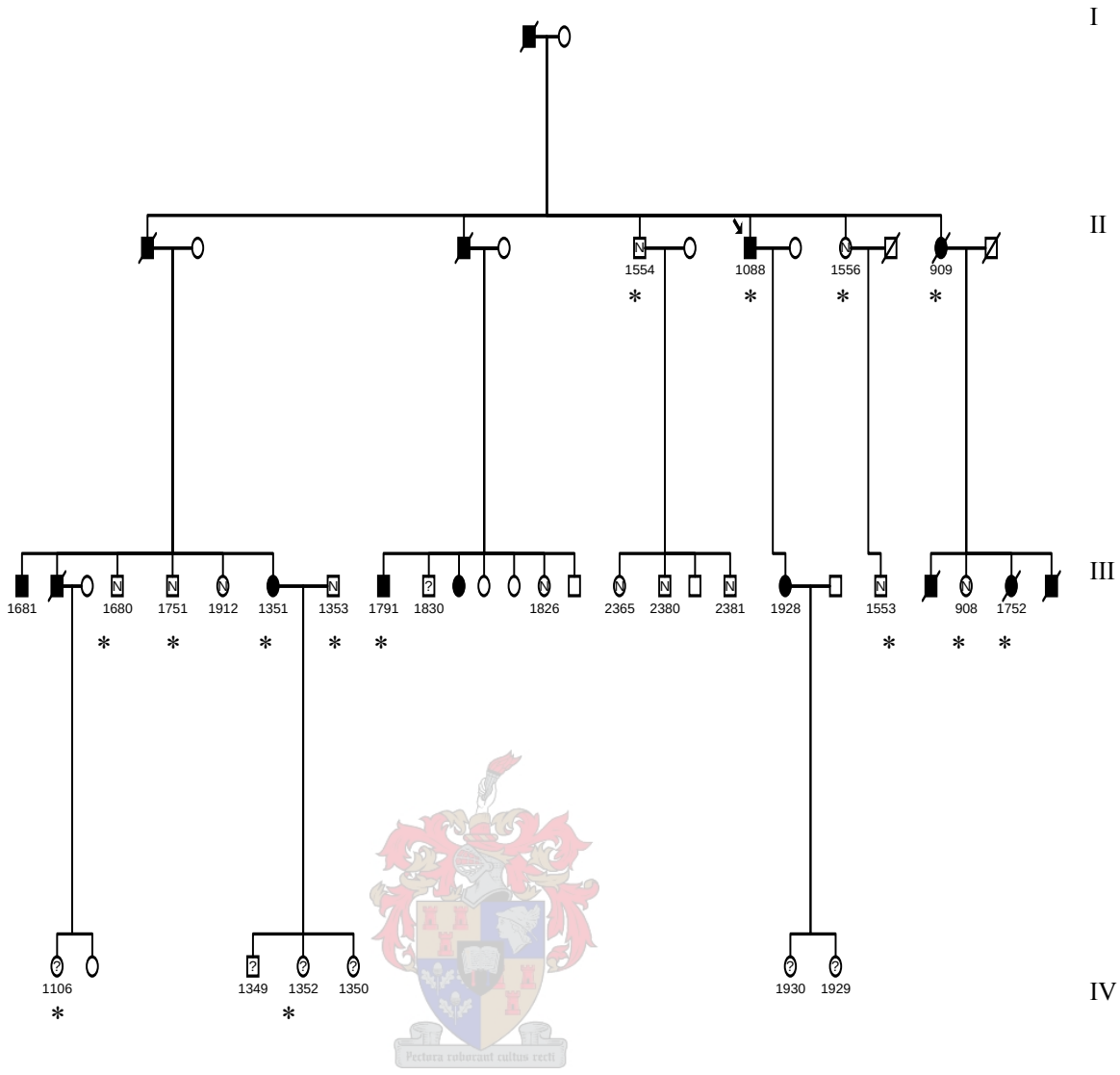


Figure 2.3 A four-generation pedigree of the PFHBII-affected family*. The roman figures represent the four generations of the family, squares indicate males; circles indicate females. Open symbols indicate individuals of unknown clinical status; N indicates unaffected individuals; solid symbols indicate PFHBII-affected subjects; diagonal lines indicate deceased individuals; question marks indicate individuals of unknown clinical status; arrow shows the proband. The numbers represent identification numbers of family members and those that were analysed (both affected and unaffected) are indicated by asterisks. Adapted from Fernandez et al., 2004.*

Table 2.1 *Follow-up clinical data of PFHBII-family members that were clinically assessed in the original study (Brink and Torrington, 1977) and screened for the PFHBII-causing mutation in the present study.*

Individual	Age (y) at diagnosis	Comments
1554	68	unaffected
1088	42	[b]SB = 48 bpm, CHB (PM 42y); present age = 69y (affected)
1556	64	unaffected
909	37	CHB (PM 37y), DCM, [+]59yrs (affcted)
1680	49	unaffected
1751	45	unaffected
1351	33	1° AVB (P-R interval = 220 m/s, March 1990), [c]SB = 38 bpm, AF, CHB (PM 33y, October 1996); present age = 38y (affected)
1353	45	unaffected
1791	41	CHB (PM 41y), DCM, [+]46y (affected)
1928	38	1° AVB (P-R interval = 240m/s, May 2000); SB = 43 bpm, Wenckebach/complete AVB (PM 38y), January 2003; present age = 38y (affected)
1553	41	unaffected
908	39	unaffected
1752	29	CHB (PM 29y), DCM, [+]43y (affected)
1106	23	chest pain, dyspnoea, syncope (atypical)

b = heart rate prior to pacemaker (PM) implant; c = individuals with atypical accompanying features

+ = age at death; 1° = first-degree; AF- atrial fibrillation; AVB = atrioventricular block; bpm = beats per minute; CHB = complete heart block; DCM = dilated cardiomyopathy; m/s = milliseconds; (PM _y) = PM age at implant; SB = sinus bradycardia, y= years (data modified from Fernandez *et al*., 2004).

2.4 Mutation analysis by polymerase chain reaction-single strand conformation polymorphism (PCR-SSCP) method

2.4.1 Principle

Single-strand conformation polymorphism is a PCR-based mutation detection technique (PCR-SSCP) that detects sequence changes, including single-base substitutions, as differences in electrophoretic mobility (Orita *et al*., 1989). The PCR-SSCP analysis is based on the principle that the primary structure of a single-stranded DNA molecule changes under mildly denaturing conditions, to a specific sequence-based secondary structure. The secondary structure is due to base pairing between nucleotides within each strand and depends on the length of the strand, the location and number of regions of base pairing, and the primary sequence of the strand. Single-stranded DNA molecules that differ by at least one nucleotide will assume different conformations, and hence, have different electrophoretic mobilities when run in non-denaturing polyacrylamide gels.

PCR-SSCP has been used to detect mutations and polymorphisms in small amplicons (100-300bp) and provides a useful and cost effective alternative for the direct detection of genetic variation, particularly when large numbers of samples are required for analysis (Cotton, 1997; Gasser, 1997).

2.4.2 Polymerase chain reaction (PCR)

Target regions of candidate genes to be analysed by PCR-SSCP were amplified using PCR, a technique used for rapidly amplifying DNA sequence exponentially by extending two short sequences (primers) that are annealed or hybridised to a unique region flanking the target DNA sequence (Saiki *et al*., 1985).

2.4.2.1 Primer design

Oligonucleotide primers that were used for amplifying the coding and promoter regions of two candidate genes, *PFKFB2* and *ATF3*, were designed from genomic sequences of these genes using DNAMAN (Lynnon Biosoft Copyright) and Primer Premier (Premier Biosoft International) computer programs. The genomic sequences of *PFKFB2* and *ATF3* are available in the NCBI nucleotide database under the following accession numbers: *PFKFB2*- AJ005577 or AJ005578, *ATF3*- L19871. Melting temperatures (tables 2.2-2.5) and the GC content of the primers were also determined using the above primer-designing programs. Primers for amplifying exons of candidate genes were designed in such a way that portions of introns preceding and following the exons of interest were included (Fig. 2.4), which allowed investigation of the intron/exon boundaries (splice sites). These target sequences will be referred to as "coding regions" or "exons" in the text. Mutations in the splice sites can affect pre-RNA splicing, the size of the gene product and possibly its function, as they result in activation of cryptic splice sites, exon skipping, creation of a pseudo-exon within an intron, or intron retention (Watkins *et al*., 1995). For amplifying larger exons (exons greater than 310 bp), two overlapping pairs of primers were designed because of the limited DNA fragment size that is suitable for detection of sequence variations on SSCP analysis (Sheffield *et al*., 1993). The primer sequences were synthesised according to the standard phosphoramide method at the Department of Biochemistry, University of Cape Town (UCT, Cape Town, South Africa), and their sequences are shown in tables 2.2-2.5.

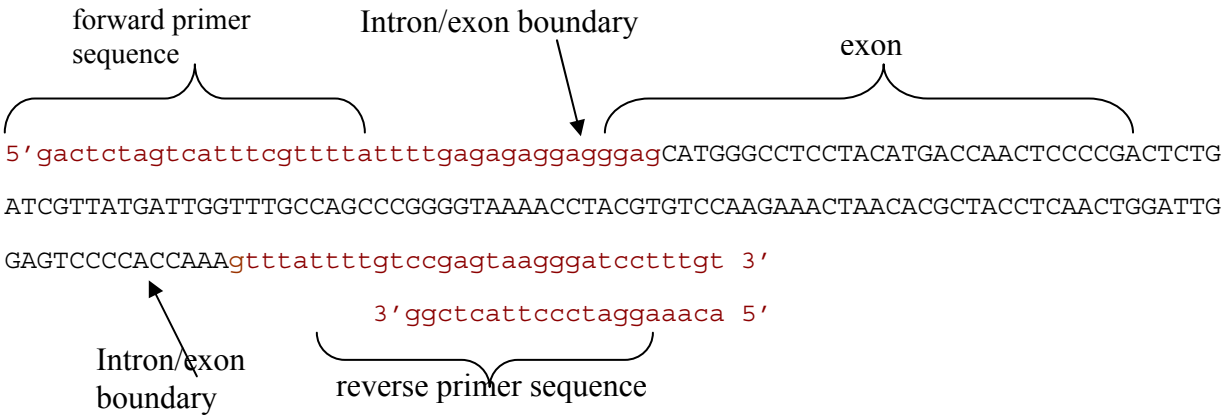


Figure 2.4 Nucleotide sequence of* PFKFB2 *intron 2-exon 3-intron 3. *Intron sequences in lower case and exon in upper case. The forward and reverse primers that were used for amplifying exon 3 are shown in colour, and intron/exon boundaries are also indicated.*

2.4.2.2 PCR conditions

DNA was amplified in a 50µl-reaction mixture containing approximately 100ng genomic DNA as a template, 75µM of each dATP, dTTP, dCTP and dGTP (Promega Corp., Madison Wisconsin USA), 5µl of a 10x Taq DNA polymerase buffer (Bioline UK Ltd, London, UK), 120ng of each oligonucleotide primer, 0.5U Taq DNA polymerase (Bioline UK Ltd, London, UK), 1.5mM $MgCl_2$, 5% formamide (Sigma Chemical Company, St. Louis, Missouri, USA) or 5% dimethylsulfoxide (DMSO) (Sigma Chemical Company, St. Louis, Missouri, USA), the latter two components being used to reduce non-specific amplification, and H_20 to a final volume of 50µl. Mineral oil was added to each tube to prevent evaporation during thermal cycling. The cycling parameters were as follows: denaturing the double stranded DNA at 94ºC for 30 seconds, annealing each primer to each single DNA strand at a specific temperature (tables 2.2-2.5) for 30 seconds, and extension of the primers by Taq DNA polymerase in the presence of dNTPs at 72ºC for 30 seconds. Amplification was performed in a thermal cycler (Gene E, Techne Ltd., Cambridge, England) and 30 cycles were done for each fragment of DNA amplified.

Table 2.2 *Oligonucleotide primers used for amplifying* PFKFB2 *coding exons, their annealing temperatures, and size of generated fragments.*

Exon	Primer sequence (5'-3')	Annealing temperature (^{o}C)	Product size (base pairs)
3	F-GACTCTAGTCATTTCGTTTTATTTTG R-TCTTTCCTAGGGAATGAGCC	48	196
4	F-GCTGCTGAAGGATTTGGGTG R-GAAAGAGCTGACCTGGAGATC	50	168
5	F-CCTTGCTGGTTTCATTTGCTC R-GACCAGAAAGAAGAAGCAGCAGA	50	130
6	F-TGCTCATTTACCTTGTGTACTTTC R-CTCATCCCTTTTCCTCCAACAT	50	144
7	F-GCTTATCACTGCCTTCTCTC R-GGAGATGATATGTAGGGGT	50	114
8	F-GAGGCCTTTACTGGTTCCT R-GGCCCTTCAAACCATGGC	50	185
9	F-TCTCAGATCTTCTTACCTAGGG R-GCCTAGAAAATCCAGTGTGCT	50	292
10	F-GTGTTTTCTTTTTACCCTTAATTTTCAT R-TCTTGTTGGCTTTTCTGGGC	50	230
11	F-CTCTTCAAGCCAGCTTCAAGTG R-AAACACTCAGCAAGGGTCTAGGC	50	190
12	F-TGCCAGCCTGCCCCATTTCC R-TGTCTCCCAGCCCCTCCCTC	50	186
13	F-GACTTCCCTCGCCTGGTCTTT R-TCCTGCCAGCCCCTATGTG	50	122
14	F-TTAAGTCCTCTCTTTCCTCTG R-GGAAACTGGTGTGTTCATCG	48	130
*15	F1-GTGGAAGCTGTTGTGGT R1-CTGACTTGGGTCTTCGGCTA	50	285
	F2-CCGACTAGCCGAAGACCCAA R2-GCCCCTGGGTGAAAAGATGT	54	225
*16	F1-ACTTCACCATTAATCTTAACACAG R1-TAGACTTTTTGTCCCCCGACT	50	298
	F2-AAGTGTTGGGGGATGGAGGG R2-TCTGGGCTGGCCCTTACAGG	57	320

F- forward primer, R- reverse primer. Asterick= overlapping primers were designed to amplify exons 15 and 16.

Table 2.3 *Oligonucleotide primers used for amplifying* PFKFB2 *promoter region, their annealing temperatures, and size of generated fragments.*

Promoter	**Primer sequence (5'-3')**	**Annealing temperature (^{o}C)**	**Product size (basepairs)**
Fragment 1	PF1-AATGAGGTTCTGCTGGGAT	50	216
	PR1-GTGGCACTGTTGCTTGGTA		
Fragment 2	PF2-GTTACCAAGCAACAGTGCCACC	56	276
	PR2-CCCTGTAACGCGACCGAA		

F- forward primer, R- reverse primer

Table 2.4 *Oligonucleotide primers used for amplifying* ATF3 *coding exons, their annealing temperatures, and size of generated fragments.*

Exon	**Primer sequence (5'-3')**	**Annealing temperature (^{o}C)**	**Product size (base pairs)**
2a	F-GGTGTTGGAGGTCTGGTG	54	238
	R-GATGGCAAACCTCAGCTCTTC		
2b	F-CACTGGTGTTTGAGGATTTTG	51	193
	R-GTGCCGAAAGAATGAATACCT		
3	F-TTCTTACTGCCCTTTAAATGTG	52	181
	R-AGAGGAAGGGTAAGGCTAGAA		
5	F-TGGCCCTTGGCTCCTTTCTT	57	266
	R-AGTCGCCCCCATACCACG		

F- forward primer, R- reverse primer

Table 2.5 *Oligonucleotide primers used for amplifying* ATF3 *promoter region, their annealing temperatures, and size of generated fragments.*

Promoter	Primer sequence (5'-3')	Annealing temperature (^{o}C)	Product size (base pairs)
Fragment 1	F-GTTCAACCTGTAAACTGGTCAAA	51	194
	R-GGGATTAGGTGGAGGAGTGT		
Fragment 2	F-CTTCATTCAAATGCAAACACTCCT	56	208
	R-CGGGTGTTTTCGCTTTACTTTTC		
Fragment 3	F-CAGGGCGAAAAGTAAAGCGAAAA	60	276
	R-TAGAGGTGTCTGGAGCCTGGGTC		
Fragment 4	F-GTGTTTCTGCCCTTCACCGTGC	57	192
	R-AAACGATTTCCTGGTTCTTATTCC		
Fragment 5	F-GGGTCGGGGAATAAGAACCAGGA	65	300
	R-GAAGAAGGTGCTTTGGGGTCCGC		
Fragment 6	F-CCTCCCACCGGGTTGCCTCTGATT	67	287
	R-TGTGACCTCTGCGGCGGGGAC		
Fragment 7	F-TAACGGAGTGGATACACCGAAGG	60	206
	R-GGCCCCAGCCCAGACTAGG		
Fragment 8	F-CTCTTCAACCTAGCGGAGGGACA	60	180
	R-CCTCGCTCACTGAGACACACACC		
Fragment 9	F-TGGGCTGGCTCCTCCCCGAACTTG	70	237
	R-CTGCGTGCGACTGTGGCTTGGAG		

F- forward primer, R- reverse primer

2.4.2.3 Agarose gel electrophoresis

In order to assess the success of non-radioactive PCR amplification, agarose gels were prepared as follows: six to eight microlitres of each PCR product was mixed with 1-2µl of bromophenol blue loading dye (appendix II) and was subsequently electrophoresed in 2% horizontal agarose gel, containing 1µg/ml ethidium bromide and 1x TBE buffer (appendix II). The λ*Pst* molecular weight marker (appendix II) was co-electrophoresed with all PCR products, which were

visualised on a longwave transilluminator (3 UV transilluminator, California, USA). Electrophoresis of agarose gels typically occurred at 10V/cm for 20 minutes in 1x TBE running buffer, and the gels, thereafter, were visualised and photographed.

2.4.3 PCR-SSCP method

2.4.3.1 Gel casting

The PCR-SSCP gels were used to detect mobility shifts in specific PCR-amplified exons of candidate genes. The gel was cast between two 390mm x 290mm glass plates, that were initially washed with Cal-liquid hand soap (Cal-Chem, R.S.A.), rinsed in tap water and dried with paper towel, and then further cleaned with 70% ethanol. The back plate was sprayed with 70% ethanol in order to fix the Gelbond™ PAG polyester film (FMC, Bioproducts, Rockland, Maine, USA) The hydrophilic side of the Gelbond™ was exposed so that it would bind covalently with the gel solution (appendix I) when poured, thereby creating a solid support for the gel for subsequent silver staining. The front plate was silinised with Wynn's C-THRU windshield rain dispersant (Wynn Oil S.A. Pty Ltd, SA). One-millilitre spacers were placed on either side of the GelBond-covered plate and the front plate was placed on top of the back plate with the silinised side on the inside. The sides of the plates and the bottom were sealed with gel-sealing tape (SIGMA, Germany) and then the casting rubber boot (S2 casting boot, Life™ Technologies, UK) was fitted for further sealing. Two types of denaturing gel solutions, 8% acrylamide + 5% glycerol and 10% acrylamide + 5%glycerol (appendix II), were prepared and poured between plates making sure that there were no bubbles that would disrupt electrophoresis of the PCR products. The glass plates were then placed down horizontally, a square-tooth, well-forming comb fitted, the plates clamped and the gel allowed to set for approximately 1 hour.

2.4.3.2 Electrophoresis

Once the gel was set, the casting boot, the sealing tape at the sides and at the bottom of the glass plates, and the comb were removed. The glass plate assembly was mounted vertically on the electrophoresis apparatus (Omeg Scientific, R.S.A), moved to the cold room (4˚C) and 0.5x TBE (appendix II) buffer added to the reservoir tanks. Before loading PCR amplified products, any residues inside the wells were removed by washing them with the electrophoresis buffer using a syringe. Five to eight microlitres of a 50µl PCR amplified product was mixed with an equal volume of SSCP loading dye (appendix II), and the samples were heat-denatured at 94˚C for 2 minutes and immediately loaded directly onto the non-denaturing gels. Control samples loaded onto the gels included a non-denatured sample as a control for electrophoretic migration of the double stranded DNA, a denatured PCR-amplified sample from an unaffected individual as a negative (wild type) control, and λ*Pst* as a molecular weight marker. Electrophoresis was carried out at a constant power of 50 watts (W) overnight at 4˚C until the dye front was three quarters of the length of the gel (depending on the product size of the amplified products).

2.4.3.3 Silver staining

Following electrophoresis of samples, the glass assembly was dismantled and the gel, still supported by the Gelbond™, was silver stained as follows: the gel was placed in a container with 50ml of solution B (0.1% AgNO3) (appendix II) and shaken gently for 10 minutes. Solution B was carefully discarded and the gel rinsed with water. Fifty millilitres of Solution C (appendix II) was then poured into the container to cover the gel, which was gently shaken until the bands were visible. The gel was then removed from the solution and hung to air-dry.

2.5 Sequencing

Mobility shifts detected by PCR-SSCP analysis indicate the presence of sequence variations that cause different alleles to migrate differently in the SSCP gel. To confirm the presence of sequence variations, and to identify a particular variation, samples with different patterns were sequenced. Before sequencing, PCR products were purified to remove all other PCR reagents and other impurities that may interfere with the sequencing reaction and yield spurious results.

2.5.1 Purification of PCR products for sequencing

PCR-amplified products to be sequenced were purified using the GFX™ PCR DNA and Gel Band Purification kit (Amersham, Bioscience, UK) following the manufacturer's instructions. The PCR-amplified products were loaded onto a 2% agarose gel, containing ethidium bromide, and electrophoresed together with the λPst molecular weight marker for approximately 1 hour in 1x TBE at room temperature. The PCR-amplified products were visualised on a longwave transilluminator and the correct size fragment, judged by the size of the molecular marker, was excised from the gel using a sterile scalpel blade. The gel was placed in a previously weighed 1.5ml eppendorf microcentrifuge tube and again weighed in order to determine the correct volume of GFX capture buffer to mix with the gel. Ten microlitres of the GFX capture buffer was added for every 10mg of the gel slice, which were then mixed vigorously and incubated at 60˚C for 10-15 minutes until the gel was completely dissolved. A GFX Microspin™ column was placed in a 2ml capless collection tube and the dissolved gel was applied to the column and incubated at room temperature for 1 minute. The sample was then centrifuged in a Beckman microcentrifuge (Germany) at full speed for 30 seconds so that the DNA could bind to the column. The flowthrough was discarded and the column replaced into the same collection tube. A GFX™ wash buffer was added into the column, which was centrifuged at full speed for 30 seconds. The collection tube was discarded and the GFX™ column was transferred to a fresh 1.5ml microcentrifuge tube. A maximum of 50µl autoclaved double distilled H_2O was added

directly to the top of the glass fibre matrix inside the GFX™ column. The sample was incubated at room temperature for 1 minute and centrifuged at full speed from 1 minute to recover the purified DNA.

2.5.2 Automated sequencing

The purified DNA fragments were sequenced using an ABI Prism™ 377 automated sequencer at the Department of Genetics (Central DNA Sequencer) at the University of Stellenbosch, Stellenbosch, South Africa.

2.5.3 DNA sequence analysis

Sequences obtained from ABI automated sequencing were compared to that of the GENBANK sequence, to ascertain sequence variations, using the NCBI pairwise2 blastn program and DNAMAN. Obtained sequences (reverse and forward sequences) were aligned using the NCBI pairwise2 BLAST program to remove ambiguous nucleotide bases and were subsequently aligned with the GENBANK sequence. The three sequences were then aligned using the Multiple Sequence alignment program in DNAMAN to identify any sequence variations and to determine their position in the gene. If there were any nucleotide changes identified, the sequence was then translated to determine the effect of the substitution on amino acid sequence. Mutation analysis was extended to a control population to determine whether the sequence variation was predicted to be disease-causing or a neutral polymorphism. In order to strengthen a functional role for the sequence variation, the analysis was extended to families (depending on the availability of DNA samples) of the individuals carrying the mutation or the polymorphism to investigate if it segregated with the disease phenotype.

2.6 Allele-specific restriction enzyme analysis

2.6.1 Identification of restriction enzyme recognition sites within fragment 5 of the ATF3 promoter region

Allele-specific restriction enzyme analysis (ASREA) was performed to confirm the presence of sequence variations that were identified and to genotype other individuals in HCM families. This was performed by determining the effect of a particular sequence variation on a specific restriction enzyme recognition sequence. The sequence analysis program, DNAMAN, was used to generate a restriction enzyme map encompassing the relevant regions of both the GENBANK and variant sequences. This analysis was performed for all sequence variations that were identified, but only two of them (in fragment 5 of *ATF3* promoter region) caused a gain or loss of the restriction enzyme recognition sequence.

Three sequence variations were identified in *ATF3* fragment 5 and two of them altered one restriction enzyme recognition sequence each. One polymorphism (-1068G/C) resulted in a loss of an *Nla*III site (CATG/) while the other (-1229C/T) abolished the *Alw*26I site (GTCTCA$(n)_1$/) that was identified in the region (Fig 2.5).

2.6.2 Restriction enzyme digestion.

The following procedure was performed on the *ATF3* promoter region in which sequence variants were identified, but using different enzymes and appropriate 10x buffers for the different variants. Five microlitres of the relevant amplified PCR-product was aliquoted into a 500µl-eppendorf microcentrifuge tube. Thereafter, 1.5µl of appropriate 10x restriction enzyme buffer (New England Biolabs), 1-2U of a restriction enzyme (*Nla*III or *Alw26*I) and water to a final volume of 15µl were added to the PCR-product. The mixture was incubated at 37˚C for 2-3 hours to allow digestion to proceed. Non-denaturing polyacrylamide gels (12%) were used for size separation of PCR-amplified products digested with restriction enzymes.

a)

```
                    EcoRII         MseI             MnlI DdeI
1      GGGGTCGGGGAATAAGAACCAGGAAATCGTTTTTAAGGTTCAAACCCAGTTCTGCTGAGG
           Alw26I      AvaII   AsuI                  BglI
       DdeI           AsuI    DraII               EcoRII   MnlI  BbvI
61     TCTCAGCTCGAATCTCGGACCACGGGGCCCCGCCTTTCCCGCCACCCTGGCTTGAGGGCA
                  Fnu4HI NlaIV HgaI                  EcoHI
121    GAGGGGATTTCTGCTGCGGGTTCCGCCTGTGGTCAGTGCGTCCCCATTCCGGGCCGTCCG
        NlaIV                        MaeII   AluI
181    GTCCCAGTCCAATCGGCTCTGGAGCGAGAGACGTGAAAGCTGAAGATGGTTTTCCCTAAA
       Eco57I       BanI                    AccII     MboII
241    TATGCCTGAGACGGCACCCCAGGCCTGGGCAGGTTCGCGGACCCCAAAGCACCTTCTTCT
```

b)

```
                    EcoRII         MseI             MnlI DdeI
1      GGGGTCGGGGAATAAGAACCAGGAAATCGTTTTTAAGGTTCAAACCCAGTTCTGCTGAGG
                     AvaII   AsuI                  BglI
      DdeI           AsuI    DraII            EcoRII   MnlI  BbvI
61     TCTTAGCTCGAATCTCGGACCACGGGGCCCCGCCTTTCCCGCCACCCTGGCTTGAGGGCA
                  Fnu4HI NlaIV HgaI                  EcoHI
121    GAGGGGATTTCTGCTGCGGGTTCCGCCTGTGGTCAGTGCGTCCCCATTCCGGGCCGTCCG
        NlaIV                        MaeII   AluI     NlaIII    MboII
181    GTCCCAGTCCAATCGGCTCTGGAGCGAGAGACGTGAAAGCTGAACATGGTTTTCCCTAAA
       Eco57I      BanI                    AccII     MboII
241    TATGCCTGAGACGGCACCCCAGGCCTGGGCAGGTTCGCGGACCCCAAAGCACCTTCTTCT
```

Figure 2.5. Restriction enzyme map encompassing a 300 bp-ATF3 promoter region*. a) GENBANK sequence, b) variant sequence. Resctriction enzyme recognition sites are printed in bold with the affected nucleotide written in blue.*

2.6.3 Gel casting and electrophoresis for detection of polymorphic restriction enzyme recognition sites

A 12% polyacrylamide gel was prepared as described in appendix II. The gel was cast between two 100mm x 80mm-glass plates (wider plates were used if many samples were to be electrophoresed) that were pre-washed and cleaned with 70% ethanol. After the gel had set, the tape was removed on both sides and the bottom of the plate, and the plates were then clamped into the gel electrophoresis apparatus and the buffer chamber filled with 1x TBE running buffer (appendix II). Bromophenol blue dye (2µl) was added to each of the digested samples (half of the restriction digest mixture). An undigested sample was co-electrophoresed, as well as the λ*Pst* molecular weight marker. Electrophoresis was done at 12 V/cm at room temperature until

the bromophenol blue had migrated the length of the gel, which was subsequently silver stained as described in section 2.5.3.3.

2.7 Statistical analysis

Genetic association studies provide a useful approach to determine the effect of DNA variants in complex diseases (Risch and Merikangas, 1996). The success of the approach depends on the accuracy of measuring the allele frequency differences between the study group and a control population. However, the problem with this approach is the likehood of obtaining spurious associations when determining small genetic effects in large population structure that consists of individuals with different ancestry (Lander and Schork, 1994). The problem can be alleviated by using family-based association studies, which use the genotypes of parents of affected individuals to ensure that association between a marker allele and the phenotype is detected only if the marker is linked to a disease locus, even in the presence of population substructure (Spielman *et al*., 1993). In the present study, a transmission disequilibrium test (QTDT) was used to determine the association (if any) between identified polymorphisms and HCM phenotypes.

2.7.1 Quantitative transmission disequilibrium tests (QTDT)

QTDT is a family based test that is used to perform linkage disequilibrium and association analysis of quantitative traits in families of any size, optionally using parental genotypes (Abecasis *et al*., 2000). It requires a pair of matched data and pedigree files as input. Data files indicate the order in which QTDT would encounter sets of data in pedigree files. A pedigree file is a set of data collected for individuals between and within families.

Set up of QTDT data

Parameters to be used in QTDT analysis were collected from the HCM files for each individual in the HCM families using Microsoft Excel. Quantitative traits that were used are as follows: interventricular septum (IVS), left ventricular mass (LVM), left ventricular posterior wall (LVPW), and left ventricular end diastolic diameter (LVED). The LVM was calculated according to the corrected American Society of Echocardiography formula: $0.8\times[1.04\times(IVS+LVED+LVPW)^3-LVED^3]+0.6$ (Devereux *et al*., 1986).

Four pedigree files were created for LVM, IVS, LVPW, LVED and IVS + LVPW. In each file, each line which described an individual included a family identifier, a personal identifier, paternal (encoded by 1) and maternal (encoded by 2) identifiers, sex code (male- 1, female 2), quantitative traits (LVM, IVS, LVPW, IVS + LVPW and LVED) and polymorphisms (encoded as two consecutive integers) that were identified (ATF3P6_839, ATF3P5_1068, ATF3P5_1229, and ATF3P3_1617). The association was determined between the quantitative traits and polymorphisms with age and gender of genotyped individuals used as covariates, because the heart size is known to be influenced by gender, body mass and age of individuals (Harshfield *et al*., 1995). Body mass could not be included because of the lack of data in the clinical files of genotyped individuals. Pedigree files created for each quantitative trait were converted into standard text files (columns separated by tabs of spaces) that can be recognised by the QTDT program (Fig. 2.6).

Running the QTDT program

HCM families that were genotyped had multiple offspring and linear models do not provide a valid test of linkage disequilibrium in such cases (Allison *et al*., 1999). Therefore, a variance components model was used which allows the test to be conducted in families that have multiple offspring and it also takes into account environment, polygenic, and additive

components (Allison *et al*., 1999; Abecasis *et al*., 2000). The model requires an addition of inheritance-by-descent (ibd) files which were obtained by running simwalk2. Simwalk2 was set by running prelude for each pedigree file. The resulting statistics were collected into a single QTDT.ibd file by running finale IBD-01.*, and then the possibility of association between the quantitative traits and polymorphisms was determined with and without covariates. The prelude and finale programs were used to convert the data in the pedigree file into a format that is compatible with simwalk2.

a) Data file without a covariate

```
T     LVM
M     P3_839
M     P5_1137
M     P5_1229
M     P6_1617
```

b) Data file with a covariate

```
T     trait
C1    covariate1
C2    covariate2
M     P3_839
M     P5_1137
M     P5_1229
M     P6_1617
```


Figure 2.6 Data files that correspond to the pedigree files that were created for the QTDT analysis*.*

a) Data file without covariates. b)When covariates are included they follow traits (T) and covariates are followed by markers (M).

CHAPTER 3

RESULTS

INDEX

3.1 Genetic analysis of HCM, DCM and cardiac conduction system disorders

The catalogue of genes previously identified as causes of HCM, DCM, or cardiac conduction system disorders is presented in tables 3.1 and 3.2. During the course of the study, complete sequences (DNA and protein sequences) of these genes were available in several databases under accession numbers that are also listed in the tables.

3.2. Identification of candidate genes within the PFHBII locus

3.2.1 BLAST hits to the PFHBII locus

Nucleotide and protein sequences of the genes in the catalogue (tables 3.1 and 3.2) were used in BLAST analyses to search for genes at the PFHBII locus (chr 1: 206860725-209674431bp) with possible homologous functional and/or structural domains to them. No BLAST hits were identified within the PFHBII locus. Homologous sequences identified on chromosome 1 were all found outside of the PFHBII locus (tables 3.3 and 3.4).

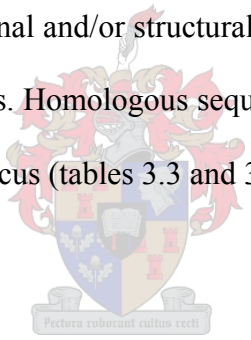


Table 3.1 *Chromosomal loci, genomic organisation, and accession numbers of sarcomeric and energy regulating protein-encoding genes implicated as causes of HCM and DCM.*

Disease	Disease gene	Locus	Genomic organisation	Accession number
	Sarcomeric proteins			
HCM and DCM	*TNNT2*	1q32	16 exons[1]	NM_000364
HCM	*TNNI3*	19q13	8 exons[2]	NM_000363
HCM and DCM	*TPM1*	15q22	9 exons[3]	M19713
HCM and DCM	*MYH7*	14q12	40 exons[4]	NM_000257
HCM	*MYL3*	3p21	7 exons[5]	NM_000258
HCM	*MYL2*	12q23-p21	7 exons[6]	S69022
HCM and DCM	*ACTC*	15q14	6 exons[7]	NM_005159
HCM and DCM	*MYBPC3*	11p11.2	35 exons[8]	Y10129
HCM	*CRP3*	11p15	5 exons[9]	NM_003476
HCM and DCM	*TTN*	2q31	363 exons[10]	NM_003319 NM_133437 NM_133432 NM_133378
	Energy-regulating protein			
HCM	*PRKAG2*	7q36	16 exons[11]	AF087875

Superscript numbers represent references for the sequences and genomic organisation: 1- Jing *et al*., 1992; 2- Bhavsar *et al*., 1996; Barton *et al*., 1999, 3- MacLeod and Gooding, 1988; 4- Jaenicke *et al*., 1990; 5- Fodor *et al*., 1989; Kurabayashi *et al*., 1988; 6- Dalla Libera *et al*., 1989; 7- Strausberg, 2001 (direct submission) 8- Carrier *et al*., 1997, 9- Arber *et al*., 1994; Jetten, 1996 (direct submission), 10 Freiburg *et al*., 2000; Cazorla *et al*., 2000; Bang et *al*., 2001, 11- Lang *et al*., 2000.

Abbreviations: *TPM1*- alpha tropomyosin gene; *ACTC*- cardic actin gene, *CRP3*- cardiac muscle LIM protein gene; *MYBPC3*- cardiac myosin binding protein-C gene, *MYH7*- cardiac myosin heavy chain gene, *MYL3*- essential myosin light chain gene, *MYL2*- regulatory myosin light chain gene; *PRKAG2*- 5'-AMP-activated protein kinase, gamma subunit gene, *TNNI3*- cardiac troponin I gene; *TNNT2*- cardiac troponin T gene.

Table 3.2 *Chromosomal loci, genomic organisation, and accession numbers of cytoskeletal, sarcoplasmic reticulum, ion channel-, nuclear- and cytoplasmic protein-encoding genes identified as causes of DCM, HCM and CCS disorder accompanied by arrhythmia.*

Disease	Disease gene	Locus	Genomic organisation	Accession number
	Cytoskeletal protein-encoding genes			
DCM	*DES*	2q35	9 exons[1]	NM_001927
DCM	*DMD*	Xp21.2	79 exons[2]	NM_004006
DCM	*SGCD*	5q33	8 exons[3]	NM_000337
DCM	*Meta-VCL*	10q22-q23	22 exons[4]	NM_014000
	SR protein-encoding gene			
DCM	*PLN*	6q22.1	2 exons[5]	NM_002667
	Nuclear protein-encoding gene			
DCM	*LMNA*	1q21.2-q21.3	12 exons[6]	NM_170707
	Ion channel-encoding gene			
CCS defects associated with arrhythmias	*SCN5A*	3p21-p24	28 exons[7]	NM_000335
	Cytoplasmic protein-encoding gene			
DCM	G4.5	Xq28	11 exons[8]	NM_181314

Superscript numbers represent references for the sequences and genomic organisation: 1- Li *et al*., 1989; Park *et al*., 2000, 2- Roberts *et al*., 1993; Koenig *et al*., 1987; 3- Nigro *et al*., 1996; Jung *et al*., 1996; 4- Koteliansky *et al*., 1992; 5- McTiernan *et al*., 1999, 6- Lin and Worman, 1993; , 7- Gellens *et al*., 1992, 8- Bione *et al*., 1996
Abbreviations:CCS- cardiac conduction system; *DES*- desmin gene; *DMD*- dystrophin gene; *G4.5*- tafazzin gene; *LMNA*- lamin A/C gene; *PLN*- phospholamban gene; *SCN5A*- voltage gated sodium channel, type 5, alpha subunit gene; *SGCD*- deltasarcoglycan gene; SR- sarcoplasmic reticulum*; VCL*- metavinculin gene.

Table 3.3 *Nucleotide sequence homologies of genes in the catalogue (tables 3.1 and 3.2) to positions on chromosome 1 and their scoring matrices. The PFHBII locus lies within chromosome 1: 206860725-209674431bp.*

Gene	accession number	Chromosome 1 locus	e-value	Percentage identity
ACTC	NM_005159	77662133 to 77662174	8.3e -127	85.71
CRP3	NM_003476	198742132 to 198753688	6.8e -52	70.58
DES	NM_001927	153301537 to 153322820	6.6e-15	55.41
DMD	NM_004006	38963395 to 39162205	2.1e -12	59.62
G4.5	NM_181314	67143243 to 67143472	3.2	59.02
LMNA	NM_170707	153301290 to 153301857	0	100.00
Meta-VCL	NM_014000	No match	-	-
MYBPC3	NM_000256	200426296 to 200432989	7.8e -79	63.54
MYH7	NM_000257	16327498 to 16334751	1.4e -07	58.96
MYL2	S69022	No match	-	-
MYL3	NM_000258	No match	-	-
PLN	NM_002667	114613339 to 114613643	1.0	57.72
PRKAG2	AF087875	149300964 to 149301256	9.6	56.25
SCN5A	NM_000335	198315815 to 198318076	1.4e -19	62.56
SGCD	NM_000337	No match	-	-
TTN	NM_133432 (novex-1 isoform)	225439641 to 225441264	3.2e-11	60.00
	NM_133378 (N2-A isoform)	183284683 to 183301618	5.4e-14	53.80
TNNI3	NM_000363	198667721 to 198672641	3.4e -49	70.80
TNNT2	AY277394	198616402 to 198616640	1.7e-140	100.00
TPM1	NM_000366	151358569 to 151381283	2.2e -112	78.86

Abbreviations: *ACTC*- actin gene, *PRKAG2*- 5'AMP-activated protein kinase, γ_2 subunit gene, *DES*- desmin gene, *DMD*- dystrophin gene, *LMNA*- lamin A/C gene, *meta-VCL*- meta-vinculin gene, MYL2- myosin essential light chain gene, *MYH7*- myosin heavy chain gene, *CRP3*- muscle LIM protein gene, *MYL3*- myosin regulatory light chain gene, *MYBPC3*- cardiac myosin binding protein-C gene, *PLN*- phospholamban gene, *SCN5A*- voltage gated sodium channel, type 5, alpha subunit gene, *SGCD*- δ-sarcoglycan gene, *G4.5*- tafazzin gene, *TNNI*- troponin I gene, *TNNT*- troponin T gene, *TTN*- titin gene, TPM1- α-tropomyosin gene. *= the full coding sequence of titin was not available at the time of the study.

Table 3.4 *BLASTP hits of genes in the catalogue (table 3.1) to positions on chromosome 1 and their scoring matrices.*

Gene	Accession number	Chromosome 1 locus	e-value	Percentage identity
ACTC	AAB59619	235131283 to 235132054	4.7e -107	79.84
AMPK-γ_2	NP_057287	36066720 to 36069724	1.5	32.63
DES	NP_001918	153317297 to 153322920	2.2e -25	28.78
DMD	NP_003997	39110150 to 39162448	8.6e -105	43.91
LMNA	NP_005563	153301502 to 153324272	1.1e-226	82.48
Meta-VCL	NP_054706	143311229 to 143315731	0.66	25.54
MELC	NP_000249	1796195 to 1796890	2.9e -13	33.10
MHC	NP_000248	16653976 to 16668348	3.6e -13	21.83
MLP	NP_003467	195177397 to 195178699	5.8e-4	34.48
MRLC	NP_000423	242192118 to 242197380	3.5e -08	31.63
cMyBP-C	NP_000247	198473136 to 198479070	1.4e -47	33.65
PLB	NP_002658	No match	-	-
SCN5A	NP_932173	178718998 to 178787734	2.1e -99	34.76
SGCD	NP_000328	No match	-	-
TAZ	NP_851831	21888484 to 21888540	4.7	47.37
cTnI	NP_000354	42456138 to 42886222	9.6	22.97
cTnT	AAP96757	198622582 to 198624248	2.9e-86	100.00
TTN	NP_597676 (novex-1 isoform)	No match	-	-
	NP_596869 (N2-A isoform)	No match	-	-
α-Tm	NP_000357	15030968 to 15049730	1.3e-3	25.11

Abbreviations: ACTC- actin, AMPK-γ_2- AMP-activated protein kinase, γ_2 subunit, DES- desmin, DMD- dystrophin,, LMNA- lamin A/C, Meta-VCL- meta-vinculin, MELC- myosin essential light chain, MHC- myosin heavy chain, MLP- muscle LIM protein, MRLC- myosin regulatory light chain, cMyBP-C- myosin binding protein-C, PLB- phospholamban, SCN5A- , SGCD- δ-sarcoglycan, TAZ- tafazzin, cTnI- troponin I, cTnT- troponin T, TTN- titin, α-Tm- α-tropomyosin

3.2.2 Genes identified using Ensembl and UCSC tools

As no genes were identified at the PFHBII locus with domains that were homologous to the genes in the catalogue, the Ensembl and UCSC gene identification tools were used. These resources allow the user to view genes residing in defined chromosomal regions and provide detailed information about these genes through linkage with other databases, information that cannot be obtained directly through similarity searches. During the initial stages of the study, the human genome sequence was in a draft phase and gene annotation was still in progress but as the study progressed, more genes were mapped to the PFHBII locus.

3.2.2.1 Ensembl gene identification tools

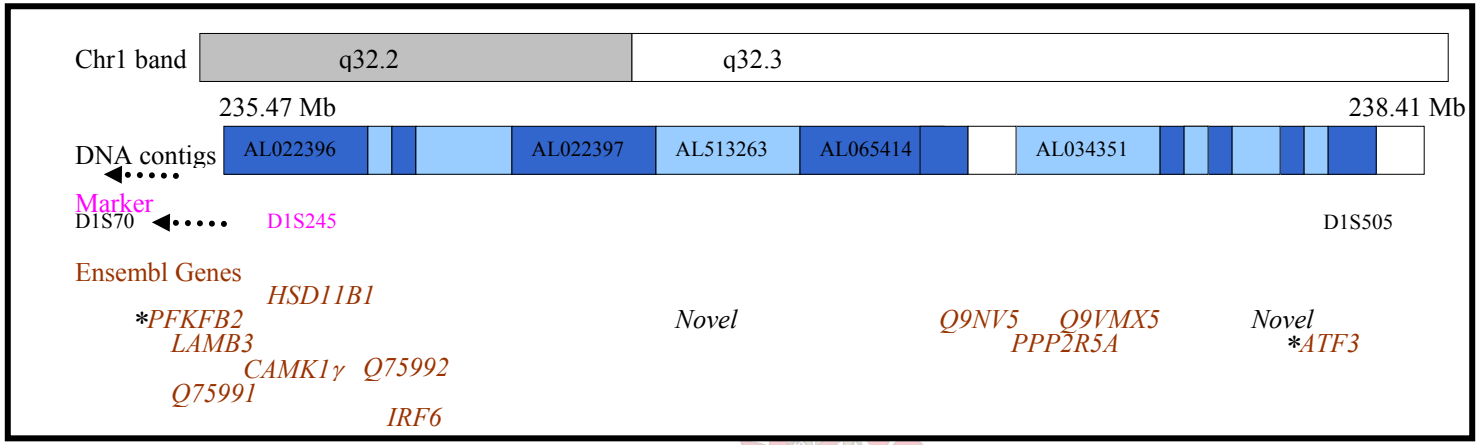
Two genes, *PFKFB2* and *ATF3*, were selected as strong candidates from the list of genes within the PFHBII locus that was obtained using the Ensembl genome browser. These genes were chosen as plausible candidates for both PFHBII-causing and HCM main locus and/or HCM modifying mutation analysis. The selection of *PFKFB2* and *ATF3* was based on the genome annotation data that was produced from the draft human genome sequence released in June 2001. At the time of the project, both genes were mapped to the PFHBII locus (Fig. 3.1a) (http://www.ensembl.org) and therefore were entered into the mutation screening protocol. However, *PFKFB2* and *ATF3* were later mapped outside the PFHBII locus due to map refinements that occurred throughout the progress of the HGP (Fig. 3.1b) (http://www.ensembl.org).

Apart from being candidate genes by position, these genes were also chosen as strong candidate genes because they are expressed in cardiac tissue and play an important role in cardiac function and/or development. *PFKFB2* encodes a 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFK-2/FBPase-2) heart isoform, which catalyses the synthesis and degradation of fructose 2,6-bisphosphate, a potent stimulator of 6-phosphofructo-1-kinase

(PFK-1), which is a key enzyme of glycolysis (Heine-Suner *et al*., 1997). Glycolysis is one of the important energy generating processes in the heart. In addition, PFK-2/FBPase-2 is a substrate for AMPK, an enzyme important in ATP regulation and in which HCM-causing mutations have been described (section 1.5.2).

ATF3 codes for activating transcription factor 3 (ATF3), which is a member of the mammalian activating transcription factor/cAMP responsive element-binding protein (ATF/CREB) family of basic region-leucine zipper (bZip) transcription factors (Hoeffler *et al*., 1988; Gonzalez *et al*., 1989; Hai *et al*., 1989; Foulkes *et al*., 1991). ATF3 homodimer is a transcriptional repressor, but its heterodimeric complex with other transcription factors (e.g., c-Jun, ATF2, JunB and gadd153/CHOP10) has been demonstrated to function as transcriptional activator (Hsu *et al*., 1992; Chu *et al*., 1994). *ATF3* is a stress-inducable gene with its mRNA level greatly increased upon cell exposure to stress signals. *ATF3* is induced in a variety of cell types by many different stress signals including myocardial ischemia-reperfusion, renal ischemia-reperfusion, chemical toxicity in the liver and partial hepatectomy, suggesting that it may be a key regulator in cellular stress responses (Chen *et al*., 1996; Yin *et al*., 1997). The common feature about the signals that induce *ATF3* is that they also induce cellular damage. Furthermore, expression of *ATF3* in cardiac tissues of transgenic mice has been reported to lead to conduction abnormalities and contractile dysfunction, suggesting that *ATF3* induction by stress signals may play a role in the pathogenesis of stress-associated cardiac diseases (Okamoto *et al*., 2001).

a)


Chr1 band
q32.2
q32.3
235.47 Mb
238.41 Mb
DNA contigs
AL022396
AL022397
AL513263
AL065414
AL034351
Marker
D1S70
D1S245
D1S505
Ensembl Genes
HSD11B1
*PFKFB2
LAMB3
CAMK1γ
Q75992
Q75991
IRF6
Novel
Q9NV5
Q9VMX5
PPP2R5A
Novel
*ATF3

b)

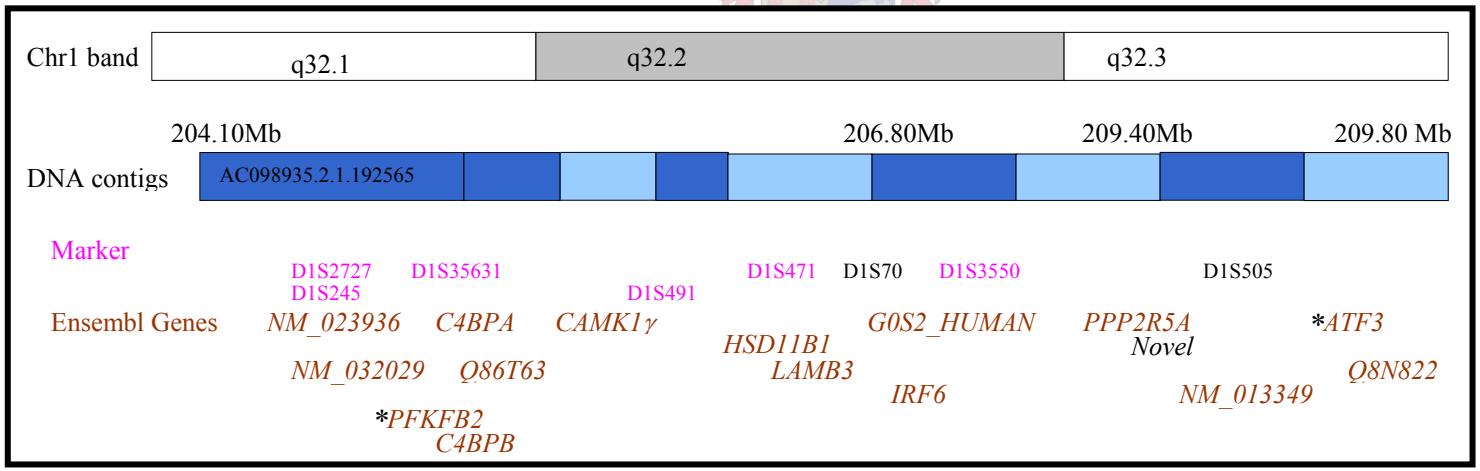

Chr1 band
q32.1
q32.2
q32.3
204.10Mb
206.80Mb
209.40Mb
209.80 Mb
DNA contigs
AC098935.2.1.192565
Marker
D1S2727
D1S35631
D1S245
D1S491
D1S471
D1S70
D1S3550
D1S505
Ensembl Genes
NM_023936
C4BPA
CAMK1γ
HSD11B1
G0S2_HUMAN
PPP2R5A
Novel
*ATF3
NM_032029
Q86T63
LAMB3
IRF6
Q8N822
NM_013349
*PFKFB2
C4BPB

Figure 3.1 Cytogenetic map of chromosome 1q32.2-1q32.3*. a) The positions of* PFKFB2 *and* ATF3 *within the PFHBII locus according to the HGP data released on June 2001. b) The positions of* PFKFB2 *and* ATF3 *outside the PFHBII locus according to the current HGP data. The PFHBII locus is defined by markers* D1S70 *and* D1S505*, which are printed in black. The chromosome 1 band (chr 1 band), DNA contigs, markers, and Ensembl known and novel genes are indicated. Abbreviations:* ATF3- *activating transcription factor 3 gene;* CAMK1γ- *calcium/calmodulin-dependent protein kinase type I;* HSD11B1- *11-beta-hydroxysteroid dehydrogenase;* LAMB3- *laminin beta-3 chain precursor;* IRF6- *interferon regulatory factor 6;* PFKFB2- *6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2;* PPP2R5A- *serine/threonine protein phosphatase 2A, alpha isoform. Adapted from the Ensembl database. Other genes in these diagrams did not have full names.*

3.2.2.1.1 Genomic organisation of candidate genes

PFKFB2

PFKFB2 consists of 16 exons that encode a protein of 505 amino acids. PFK-2/FBPase-2 is a monomer divided into two functional domains, namely, the kinase and the phosphatase domains, with the kinase of the heart isoform regulated by the C-terminal regulatory domain (Kurland *et al*., 1992; Kutamura and Uyeda, 1988*).* Exons 3-8 encode the kinase domain and exons 9-14 the phosphatase domain and the regulatory domain is encoded by exon 15 (Fig. 3.2) (Heine-Suner *et al*., 1998).

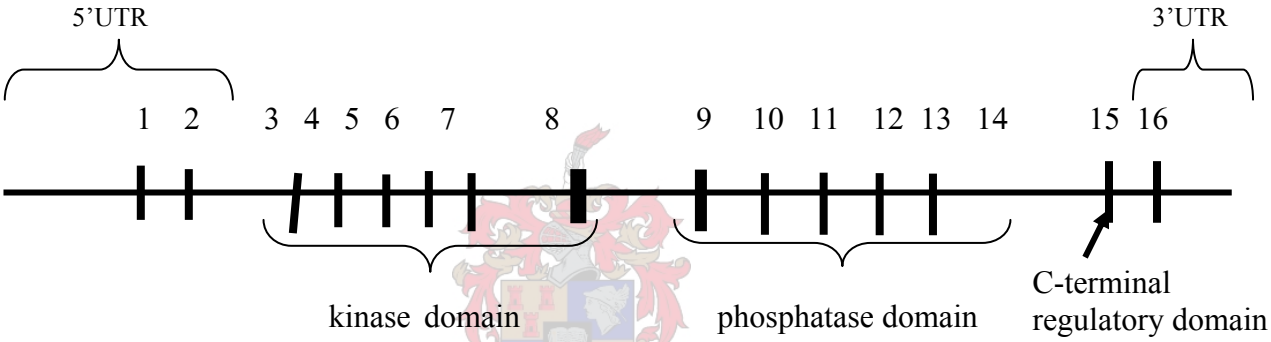


***Figure 3.2 Genomic organisation of* PFKFB2**. *Bars and numbers represent exons. The functional domains are also indicated.*

ATF3

ATF3 is encoded by four exons (A, B, C, and E) which are distributed over 15 kilobases (Liang *et al*., 1996) (Fig 3.3). The protein consists of three domains: the N-terminal domain, basic region, and the leucine zipper (ZIP) domain. ATF3 has an alternatively spliced isoform, ATF3ΔZip (Chen *et al.,* 1994) which has an additional exon (exon D) that introduces an in-frame termination codon resulting in a truncated protein (Fig. 3.3).

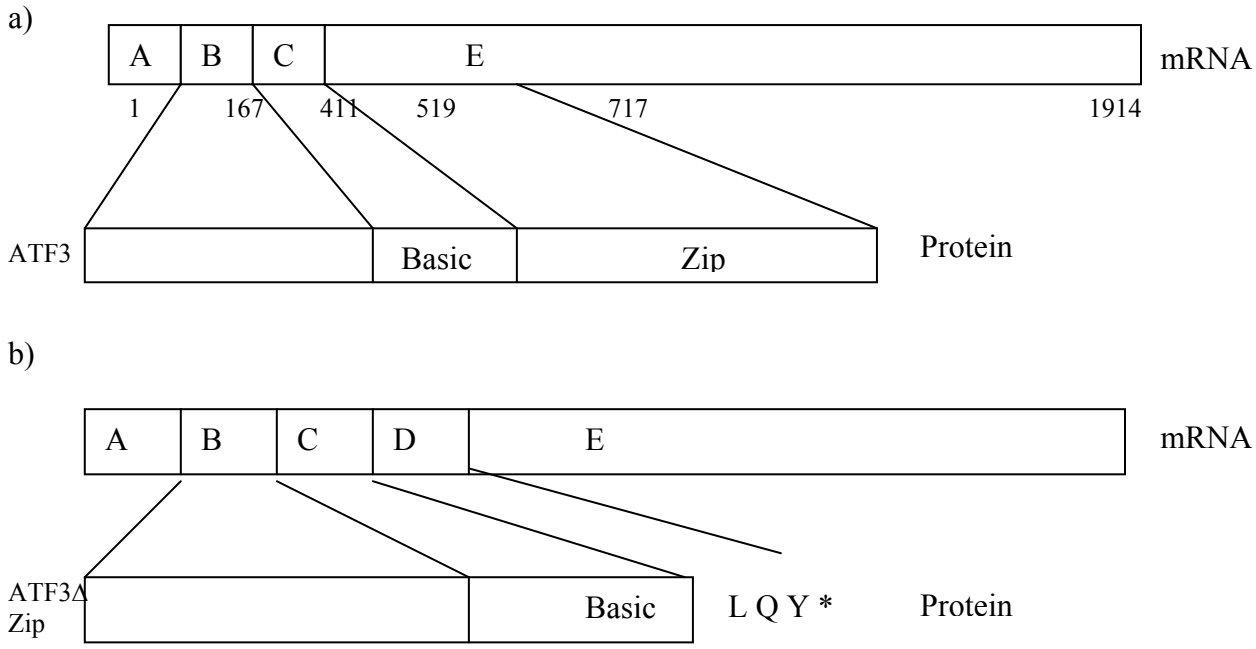


***Figure 3.3 Genomic organisation of a)* ATF3 *and b)* ATF3ΔZip**. *Exons are represented by boxes labelled A, B, C, D, and E. Functional domains are indicated by boxes, the basic region and leucine zipper are labelled. The amino acids at the border are indicated. Abbreviation: Q- glutamine,Y- tyrosine, *= stop codon. Taken from Liang et al., 1996.*

3.2.2.2 UCSC gene identication tools

Other strong candidate genes were identified using UCSC gene identification tools after the completion of the HGP, which was announced on 14 April 2003 (Fig. 3.5) (Collins *et al*., 2003). Apart from being candidate genes by position, these genes were chosen because they are expressed in cardiac tissue and are functionally related to other genes that were reported to be important in heart function. The functions of these genes include apoptosis regulation, calcium binding, energy regulation (ATP transportation and phospholipid metabolism), signal transduction, and transcription regulation. Any protein defects that may interfere with these functions have been reported to disrupt cardiac function (discussed in chapter 1). The quality of the sequences of the genes was also verified so that they can be entered into future studies and screened for PFHBII-causing and HCM main locus and/or modifying mutations.

Table 3.5 *Strong candidate genes identified using UCSC gene prediction tools, their sequence status and predicted functions.*

Candidate gene	Sequence status	Predicted function
G0S2	Complete	Apoptosis regulator activity
**FLJ25078*	Complete	Calcium-binding mitochondrial carrier protein, ATP dependent transporter, nucleic acid binding
IRF6	Complete	Transcription factor
**FLJ10876*	Complete	DNA binding, nuclear protein
TRAF5	Complete	Signal transduction
**KIAA0205*	Complete	Member of 10 acyl-SN-glycerol-3-phosphate acyltransferase family, phospholipid metabolism
SLC30A1/ZNT1	Complete	Zinc transporter
NIMA-NEK2	Complete	ATP-binding, regulation of mitosis
RAMP zinc finger protein 198	Complete	Regulation of transcription

Abbreviations: *IRF6*- Interferon-receptor factor 6 gene; *G0S2*- putative lymphocyte G0/G1 switch gene; *NIMA-NEK2*- Never in mitosis gene a-related kinase 2, *TRAF5*- Tumor necrosis factor receptor-associated factor 5 gene; *SLC30A1/ZNT1*- solute carrier family 30 (zinc transporter), member 1 gene. Asterisks indicate predicted genes with no full names.

3.2.3 Identification of promoter elements in the 5' region of PFKFB2 *and* ATF3

3.2.3.1 PFKFB2 *5' region*

The alignment of the 5' region of the *H. sapiens PFKFB2* with the *R. norvegicus pfkfb2* promoter region identified the positions of five reported or known transcription-binding sites (three *Sp1*- and two *HNF1*-binding sites) within a 500-bp *PFKFB2* promoter region (Fig 3.4) (Heine-Suner *et al*., 1998). The positions of these transcription-binding sites were identified so that primers for amplification of this region could be designed and to enable analysis of PFHBII-causative and HCM main locus and/or modifying mutations.

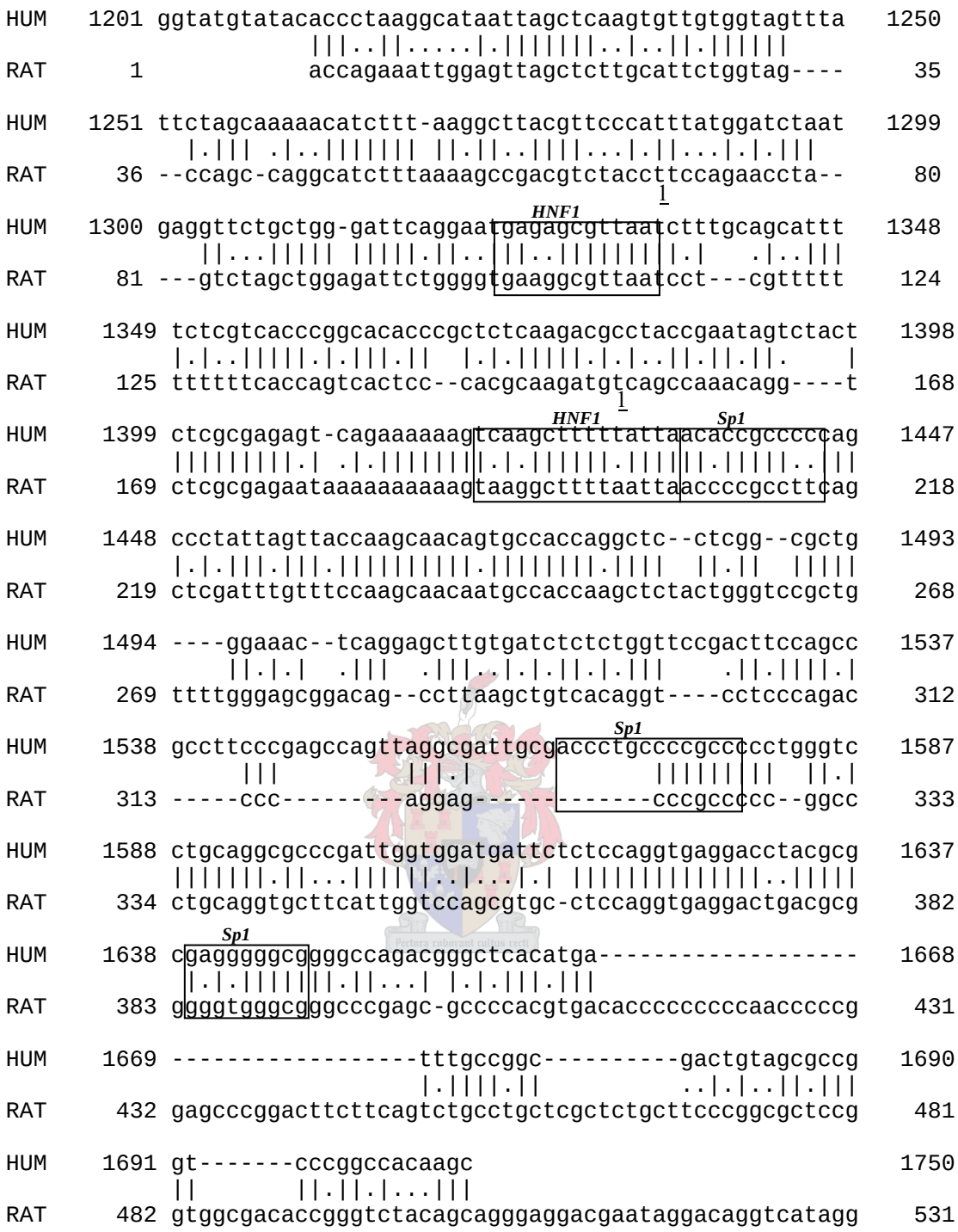


Figure 3.4 Alignment of the 5' flanking region of* Homo sapiens PFKFB2 *gene and exon 1c of* Ratus norvegicus pfkfb2 *gene. *Conserved sequences that are similar to HNF1 and Sp1 consensus sites are shown in boxes.*

3.3 Mutation analysis of candidate genes

3.3.1 Search for the PFHBII-causing mutation in the PFHBII family

3.3.1.1 PFKFB2

The coding exons of *PFKFB2* (Fig 3.2) were analysed for the PFHBII-causing mutation using PCR-SSCP analysis. Two different electrophoretic patterns were identified in exon 8 in both unaffected and affected individuals (Fig. 3.5). No mobility shifts were detected in other exons. Two samples exhibiting different patterns were sequenced but no sequence variations were identified. Samples were re-analysed using the same methods (PCR-SSCP and automated sequencing) and similar results were obtained. The 500 bp-promoter region of *PFKFB2* (Fig. 3.4) was also analysed for possible PFHBII-causing mutation but no sequence variations were identified (Fig. 3.6).

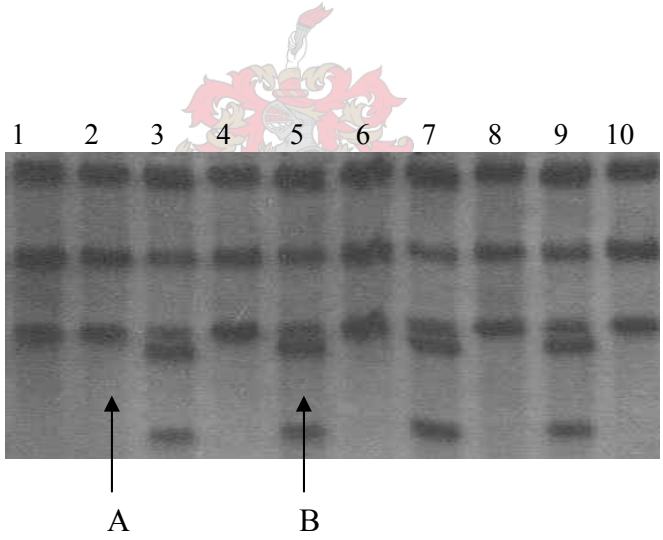


Figure 3.5 Representative PCR-SSCP analysis of exon 8 of* PFKFB2 *in the PFHBII family*. Lanes 1-3 = unaffected individuals, lanes 4-6 = affected individuals, lanes 7-10 = unaffected individuals from the general Afrikaner population. The arrows indicate two patterns, A and B, that were identified in the exon. No sequence variation could be identified to account for the different patterns detected in this exon.*

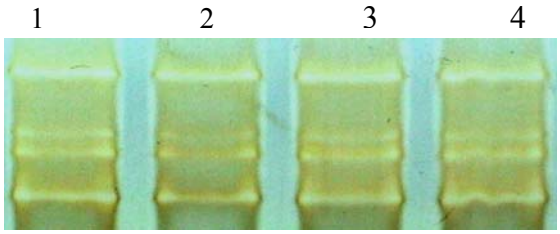


Figure 3.6 Representative PCR-SSCP analysis of fragment 2 of the* PFKFB2 *promoter region in the PFHBII family. *Lanes 1 and 2= unaffected individual, lanes 3 and 4= affected individuals. No mobility shifts were identified in this fragment.*

3.3.1.2 ATF3

The *ATF3* promoter and coding regions were analysed for the PFHBII-causative mutation in the PFHBII family but no sequence variations were identified in this gene. (Figs. 3.7 and 3.8).

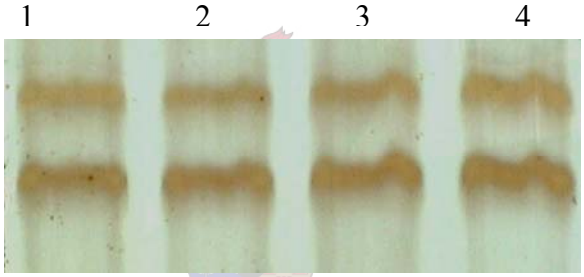


***Figure 3.7 Representative mutation analysis of exon E of* ATF3**. *Lanes 1-2= unaffected individuals, lanes 3-4= affected individuals. No mobility shifts were identified in this exon.*

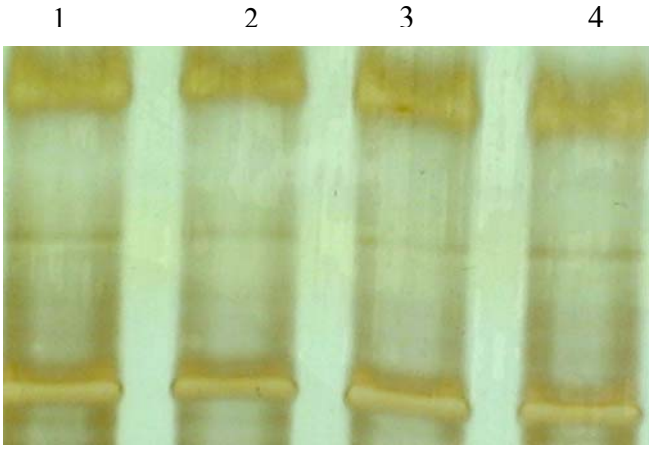


Figure 3.8 Representative mutation analysis of fragment 3 of the* ATF3 *promoter region. *Lanes 1-2= unaffected individuals, lanes 3-4= affected individuals. No mobility shifts were identified in this fragment.*

3.3.2 Search for HCM-causing and/or modifying mutations in the HCM panel

3.3.2.1 PFKFB2

Coding region

The coding region of *PFKFB2* was screened for HCM-causing and/or modifying mutations using PCR-SSCP analysis in a panel of 135 HCM-affected individuals. Different electrophoetic mobility patterns were only identified in exons 5 and 8. Three different electrophoretic mobility patterns were identified in exon 5 (Fig. 3.9). One pattern was found in individual 445 (lane 5 in figure 3.9), the second pattern was found in four of the individuals analysed (three of these individuals, 3352, 2712 and 1508 are indicated in lanes 1-3), and the third pattern (lane 4) was found in the remaining 129 individuals of the HCM panel. The three amplicons with these different patterns were subsequently sequenced, and the obtained sequences were compared with the GENBANK sequence (accession number: AJ005577) to determine the presence or absence of sequence variations. A G→A transition substitution was identified in a sample sequence of individual 445 (Fig. 3.10) at position 11535 of the genomic sequence of *PFKFB2*. However, the substitution did not result in an amino acid change. This substitution is similar to the one deposited by other investigators in the NCBI-SNP database under the SNP ID number rs2629639. No sequence variation could be identified in the amplicon of individual 2712 to account for the variant pattern identified in lanes 1-3. The sequence of the amplicon representing the third pattern was similar to that of exon 5 of the published sequence.

Two patterns similar to those identified in the PFHBII family (Fig. 3.5) were also identified in exon 8 in the HCM panel (data not shown) but again no sequence variations were identified.

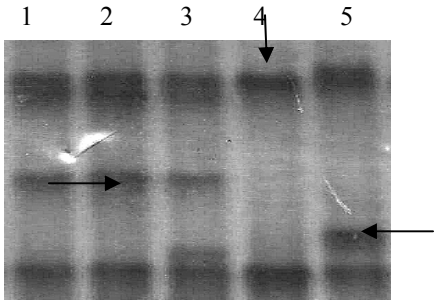


Figure 3.9 Representative PCR-SSCP analysis of exon 5 of* PFKFB2 *in the HCM panel. *The arrows indicate three different patterns that were identified in the HCM panel. Lane 1=3352, lane 2= 2712, lane 3= 1508, lane 4= 2557 and lane 5= 445. The variant pattern caused by the G→A transition substitution identified at position 11535 is indicated in lane 5. No sequence variation could be identified to account for the variant pattern identified in lanes 1-3. No sequence variant was identified in the sequence of individual 2557 represented in lane 4*

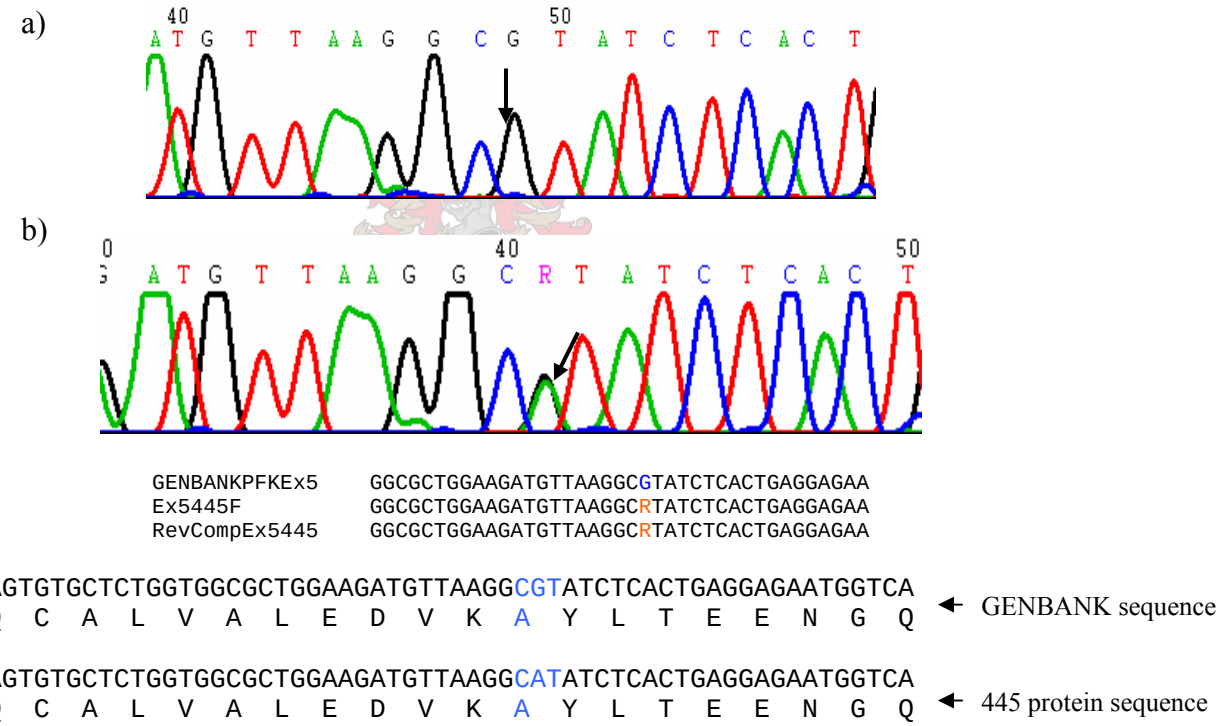


Figure 3.10. Representative mutation analysis of exon 5 of* PFKFB2 *in individual 445. *Chromatogram representing the partial nucleotide sequence of a) individual 2557 (*GG *genotype) and b) individual 445 (G/A variant) with the position of the sequence variation indicated by an arrow. Below is the CLUSTAL aligment of the GENBANK PFKFB2 sequence and the forward and reverse complementary sequences of individual 445. Translation of the two sequences is also shown with the affected codon written in colour. R refers to* G/A *heterozygote.Abbreviations: NCBIPFKEx5- exon 5 of the GENBANK sequence, Ex5445F- forward sequence of* PFKFB2 *exon 5 of individual 445, RevComplEx5445- reverse complementary sequence of exon 5 of individual 445*

Promoter region

The *PFKFB2* promoter region was also analysed for HCM main locus and/or modifying mutations but no sequence variations were identified (Fig. 3.11).

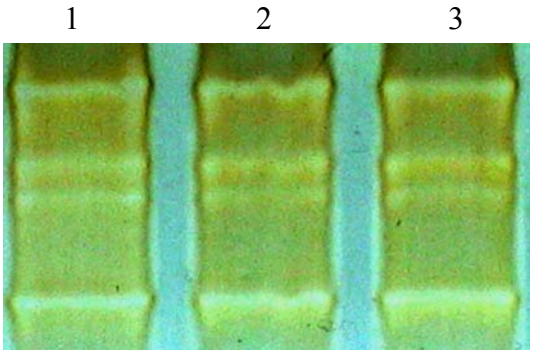


Figure 3.11 Representative PCR-SSCP analysis of fragment 2 of the* PFKFB2 *promoter region in the HCM panel. *Lanes 1-3 represent some of the HCM-affected individuals in the panel that were analysed. No mobility shifts were identified in this region.*

3.3.2.2 ATF3

Coding Region

No mobility shifts were identified in the coding exons (exons B, C, and E) of *ATF3* (Fig. 3.3) analysed in the HCM panel (Fig. 3.12).

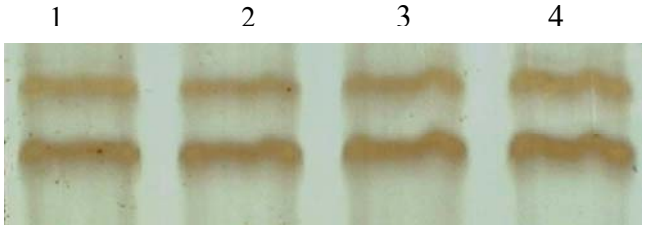


Figure 3.12 Representative mutation analysis of exon E of* ATF3 *in the HCM panel. *Lanes 1-2= unaffected individuals, lanes 3-4= affected individuals. No mobility shift were identified in this exon as indicated in the diagram*

Promoter Region

The *ATF3* promoter region was also analysed for HCM-causing and/or modifying mutations in the HCM panel. Mobility shifts were identified in four of the fragments (fragments 3, 5, 6, and 8).

Fragment 3

Two different electrophoretic mobility patterns were identified in fragment 3 of *ATF3* promoter region in the HCM panel: 14 individuals demonstrated the pattern shown in lane 2 of figure 3.13 while the remaining 133 had a pattern that is shown in lane 1 of figure 3.13. Two analysed DNA samples representing these different patterns were subsequently sequenced and the obtained sequences were aligned with the GENBANK sequence (accession number: U37542). At position -1688 a C was identified in all three amplicons that were sequenced not a T, as recorded in the GENBANK sequence (Fig. 3.14,b2). A second transition substitution (-1617A/G) (Fig. 3.14, b1) was identified in one of the obtained sequences represented in lane 2 of figure 3.13. Fragment 3 of the *ATF3* promoter region was analysed further in families of the probands in the HCM panel to determine the segregation of these sequence variants using both PCR-SSCP analysis and ASREA (Figs. 3.20-3.22), and similar variant patterns were identified.

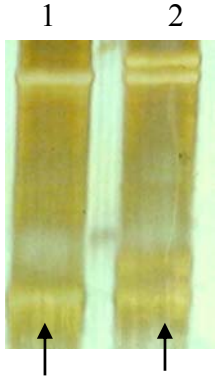
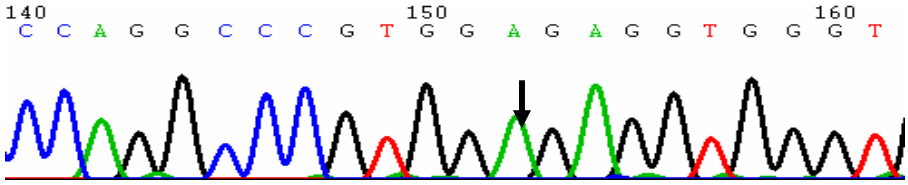
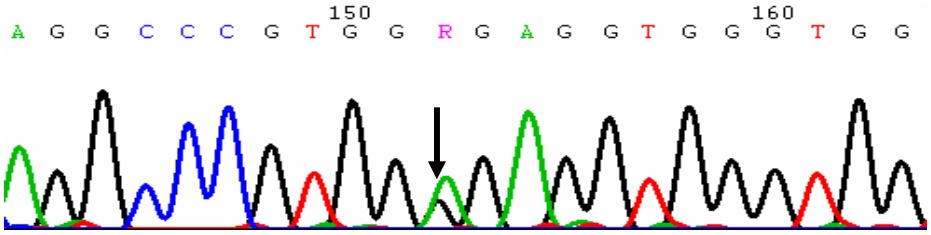


Figure 3.13 Representative PCR-SSCP analyses of* ATF3 *promoter region, fragment 3 in the HCM panel*. The arrows indicate the two different electrophoretic patterns identified in the HCM panel. Lanes 1 = individual with -1617*A/A *genotype, lanes 2 = individual with -1617*A/G *genotype (see figure3.15).*

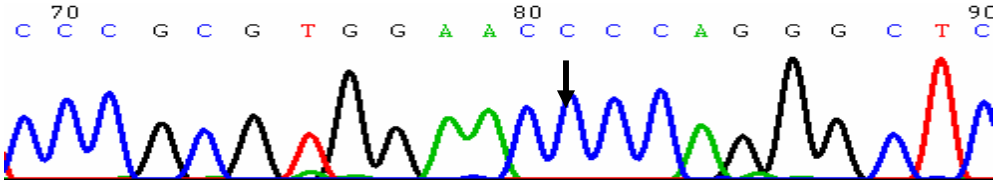
a)



b1)



b2)




```
GENBANK seq        CGCGTGGAACTCCAGGGCTCCCGGGTCCGCCGGGGCGCAAAGACTTCCGAGGCCGCCCTC
p32829F            CGCGTGGAACCCCAGGGCTCCCGGGTCCGCCGGGGCGCAAAGACTTCCGAGGCCGCCCTC
p32829Rev_Compl    CGCGTGGAACCCCAGGGCTCCCGGGTCCGCCGGGGCGCAAAGACTTCCGAGGCCGCCCTC

published atf3     CGCGTGTTCCCAGGCCCGTGGAGAGGTGGGTGGTCTGAGTGAGGTCGGGCTTGGCGGCGA
p32829F            CGCGTGTTCCCAGGCCCGTGGRGAGGTGGGTGGTCTGAGTGAGGTCGGGCTTGGCGGCGA
p32829Rev_Compl    CGCGTGTTCCCAGGCCCGTGGRGAGGTGGGTGGTCTGAGTGAGGTCGGGCTTGGCGGCGA
```

Figure 3.14 Representative mutation analysis of the* ATF3 *promoter region, fragment 3 in the HCM panel. *The chromatographs represent the partial sequences of two individuals in the HCM panel with variant patterns. a) the 1617*A/A *genotype, b1) -1617*A/G *genotype, and b2) -1688*C/C *genotype. Below the chromatogram is the alignment of the GENBANK sequence, and the forward and reverse complementary sequences obtained from automated sequencing. R refers to the* A/G *heterozygote. Abbreviations: seq- sequence, p32829F- forward sequence of an amplicon of individual 2829 in lane 2, p32829Rev_Compl- reverse complementary sequence of an amplicon of individual 2829.*

Fragment 5

Three different patterns were identified in individuals in fragment 5 of *ATF3* promoter region in the HCM panel (lanes 1, 2, and 3 in figure 3.15) and three amplicons with different patterns were subsequently sequenced. The obtained sequences were aligned with GENBANK sequence (accession number: U37542) to determine either the presence or absence of sequence variations. A G→C transversion substitution was identified at position –1068 (-

1068G/C) (Fig. 3.16) in all three sequenced amplicons. A similar SNP was deposited by other authors in the NCBI-SNP database and is available under the SNP-ID number rs4951627. The second variant, a C→T transition substitution, was identified at position –1229 of the *ATF3* promoter region (-1229C/T) (Fig 3.16) in one of the sequenced amplicons (lane 3). No additional sequence variations were identified in the third amplicon (lane 2) (Fig 3.17), though it appeared different from other amplicons in the SSCP gel. Fragment 5 of the *ATF3* promoter region was analysed further in families of the probands in the HCM panel to determine the segregation of the identified variants using both the PCR-SSCP analysis and ASREA (Figs. 3.20-3.22) and similar electrophoretic patterns were identified.

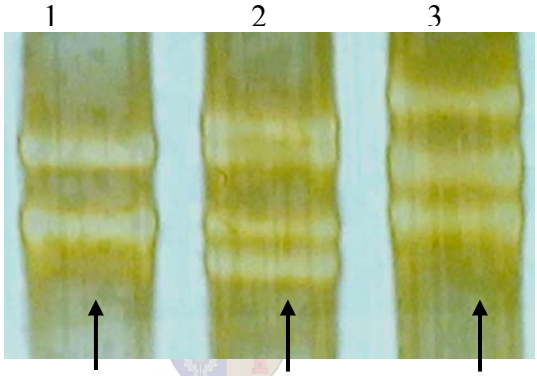
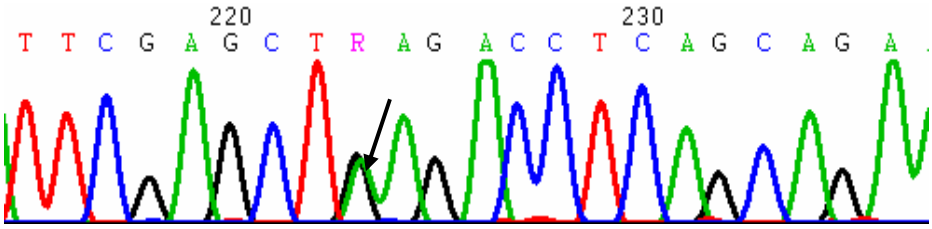
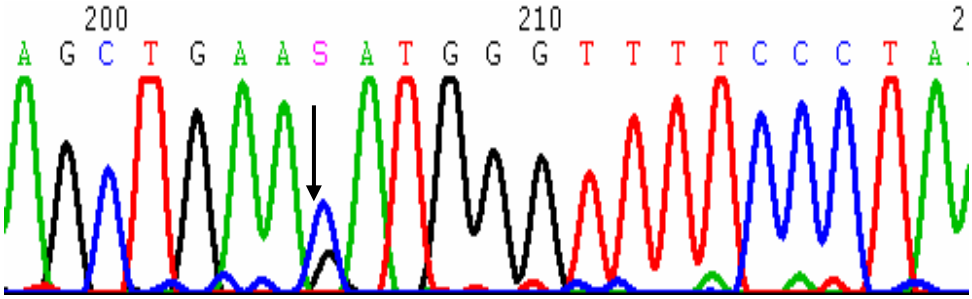


Figure 3.15 Representative PCR-SSCP analysis of* ATF3 *promoter region, fragment 5 in the HCM panel. *The arrows indicate the three patterns identified in the HCM panel. Lanes 1 and 2 = individuals with the -1229*C/C *genotype, lanes 3= individuals with the -1229*C/T *genotype (see section 3.4.2). No additional sequence variation was identified to account for the third pattern in lane 3.*

a)



b)




```
GENBANK seq          AAGAACCAGGAAATCGTTTTTAAGGTTCAAACCCAGTTCTGCTGAGGTCTCAGCTCGAA
p51031F                        ANNNNNTTTTTAGGTNNAACCCAGTTCTGCTGAGGTCTRAGCTCGAA
p51031 Rev_Compl   TAAGAACCAGGAAATCGTTTTTAAGGTTCAAACCCAGTTCTGCTGAGGTCTRAGCTCGAA

publ.atf3promoter TCTCGGACCACGGGGCCCCGCCTTTCCCGCCACCCTGGCTTGAGGGCAGAGGGGATTTCT
p51031F           TCTCGGACCACGGGGCCCCGCCTTTCCCGCCACCCTGGCTTGAGGGCAGAGGGGATTTCT
p51031 Rev_Compl    TCTCGGACCACGGGGCCCCGCCTTTCCCGCCACCCTGGCTTGAGGGCAGAGGGGATTTCT

publ.atf3promoter GCTGCGGGTTCCGCCTGTGGTCAGTGCGTCCCCATTCCGGGCCGTCCGGTCCCAGTCCAA
p51031F           GCTGCGGGTTCCGCCTGTGGTCAGTGCGTCCCCATTCCGGGCCGTCCGGTCCCAGTCCAA
p51031             GCTGCGGGTTCCGCCTGTGGTCAGTGCGTCCCCATTCCGGGCCGTCCGGTCCCAGTCCAA

publ.atf3promoter  TCGGCTCTGGGAGCAGAAGAACACGTGAAAGCTGAAGATGGGTTTTCCCTAAATATTGCC
p51031F            TCGGCTCTGGGAGCAGAAGAACACGTGAAAGCTGAASATGGGTTTTCCCTAAATATTGCC
p51031 Rev_Compl   TCGGCTCTGGGAGCAGAAGAACACGTGAAAGCTGAASATGGGTTTTCCCTAAATATTGCC
```

Figure 3.16 Representative mutation analysis of the* ATF3 *promoter region, fragment 5 in the HCM panel. *The chromatographs represent the partial sequences of two individuals in the HCM panel with different electrophoretic patterns. a) a chromatogram of a reverse sequence of fragment 5 of* ATF3 *promoter region showing the -1229*C/T *variant b) a chromatogram of a reverse sequence of the same fragment showing the -1068*G/C *variant (indicated by arrows). Below the chromatogram is the alignment of the GENBANK sequence, and the forward and reverse complementary sequences obtained from ABI traces. R refers to the* C/T *heterozygote, S refers to the* G/C *heterozygote. Abbreviations: seq- sequence, p51031F- forward sequence of an amplicon of individual 1031 in lane 3, p51031Rev_Compl- reverse complementary sequence of an amplicon of individual 1031.*

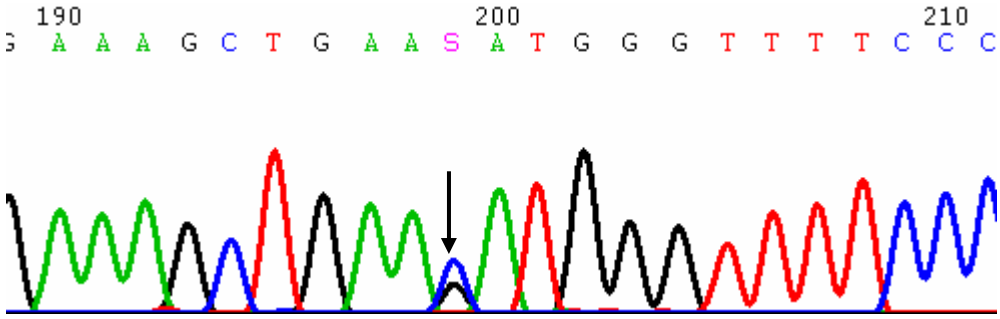
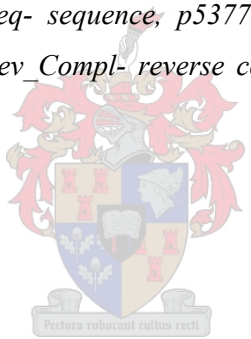


Figure 3.17 Representative mutation analysis of* ATF3 *promoter region, fragment 5 in the HCM panel*. The chromatograph represents the partial sequences of the third amplicon (lane 3 in figure 3.16).The -1068*G/C *substitution is indicated by an arrow. Below the chromatogram is the alignment of the GENBANK sequence, and the forward and reverse complementary sequences obtained from automated sequencing. The sequence variation is written in colour. No additional sequence variation was identified although this amplicon in SSCP gel appeared different to other patterns. S refers to the* G/C *heterozygote. Abbreviations: seq- sequence, p53776F- forward sequence of an amplicon of individual 1031 in lane 2, p53776Rev_Compl- reverse complementary sequence of an amplicon of individual 3776.*



Fragment 6

Two different electrophoretic patterns were identified in fragment 6 of the *ATF3* promoter region in individuals in the HCM panel: one pattern is shown in lane 1 and another pattern in lane 2 (Fig. 3.18). An amplicon with each pattern was subsequently sequenced and the obtained sequences were aligned with the GENBANK sequence. A T→A transversion substitution was identified at position –839 of the *ATF3* promoter region (-839T/A) (Fig. 3.19). Fragment 6 of the *ATF3* promoter region was analysed further in families of the probands in the HCM panel to determine the segregation of the variant (Figs. 3.20-3.22), and similar variant patterns were identified.

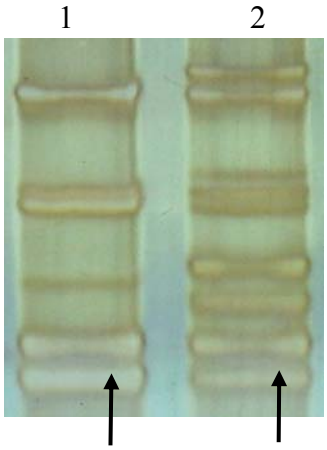
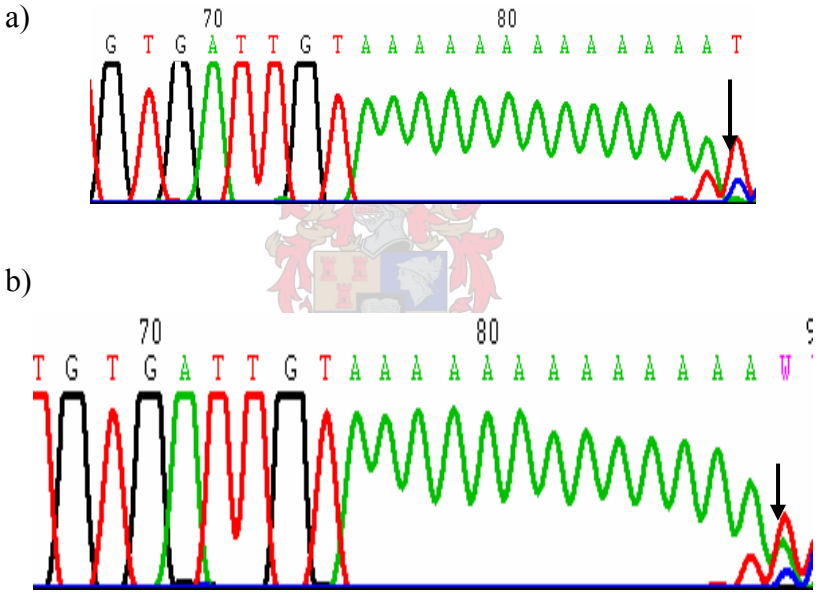


Figure 3.18 Representative PCR-SSCP analysis of ATF3 promoter region, fragment 6 in the HCM panel*. the arrows indicate two variant patterns identified in individuals in the HCM panel. Lane 1= individuals with -839T/T genotype, lane 2 = -839T/A genotype.*




```
GENBANK seq     CTTGTCAATTTCAAACGCTTTGTGATTGTAAAAAAAAAAAAAATCGAACCGATACGGTCCT
Rev Comp        CTTGTCAATTTCAAACGCTTTGTGATTGTAAAAAAAAAAAAAAWCGAACCGATACGGTCCT
P6Forward       CTTGTCAATTTCAAACGCTTTGTGATTGTAAAAAAAAAAAAAAWCGAACCGATACGGGCCT
```

Figure 3.19 Representative mutation analysis of* ATF3 *promoter region, fragment 6 in the HCM panel. *The chromatographs represent the partial sequences of two amplicons with different electrophoretic patterns. a) -839*T/T *genotype, b) -839*T/A *genotype. The position of the sequence variant is indicated by an arrow. Below the chromatogram is the alignment of the GENBANK sequence, and the forward and reverse complementary sequences obtained from automated sequencing. The sequence variation is written in colour. W refers to the* T/A *heterozygote. Abbreviations: Rev Comp- reverse complementary, seq- sequence*

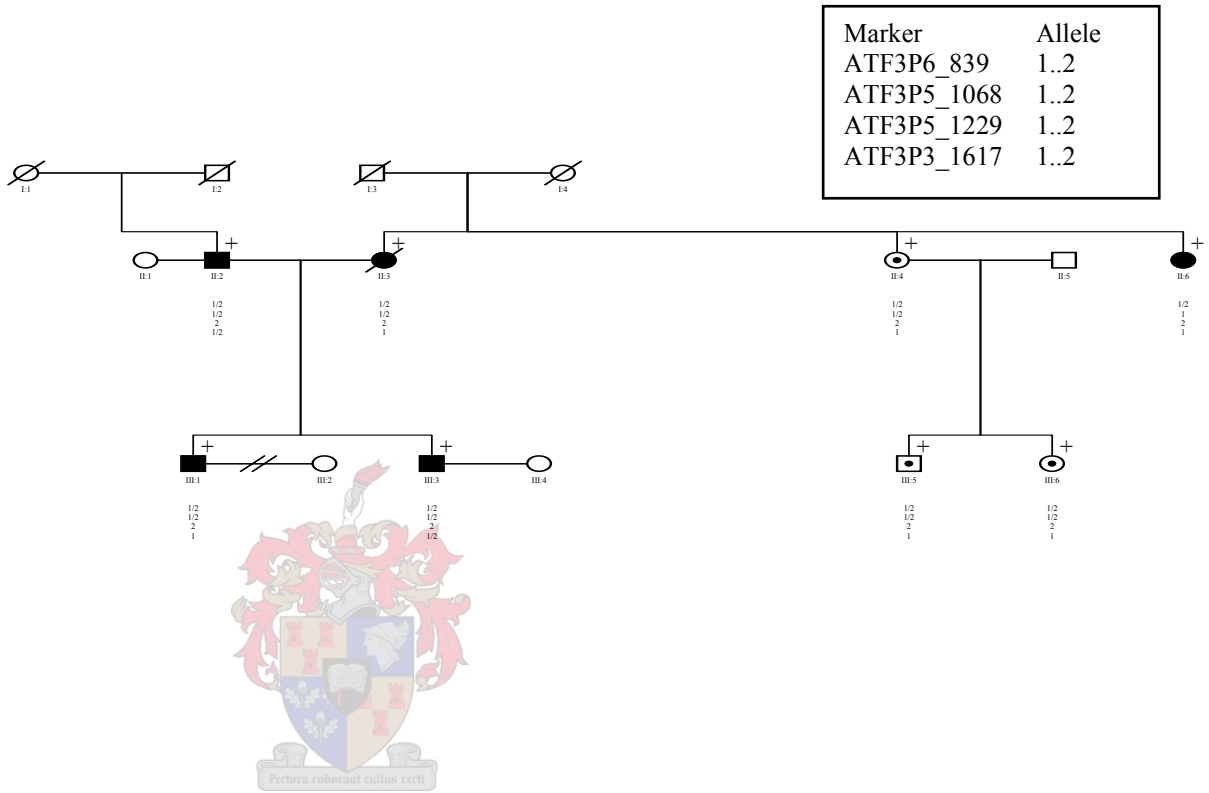


Figure 3.20 Pedigree 101, which was investigated for the segregation of single nucleotide changes identified in fragments 3, 5 and 6 of the ATF3 promoter region. *The family was reported in previous studies as carrying the* MYH7-Ala797Thr *mutation (Moolman et al., 1998; Moolman et al., 2000). Allele definition: ATF3P6_839- 1=T, 2=A; ATF3P5_1068- 1=G, 2=C; ATF3P5_1229- 1=C, 2=T; ATF3P3_1617- 1=A, G=2. 2= 2/2 and 1= 1/1. Symbol definition: Open symbols indicate individuals with unknown clinical status; solid symbols indicate affected subjects; plus symbol indicate* MYH7-Ala797hr *mutation carriers, open symbol with a dot indicates* MYH7-Ala797hr *mutation carriers but not clinically affected; diagonal lines indicate deceased individuals. The order of the marker genotypes, from the telomere to the centromere, is indicated in the box at the top of the diagram.*

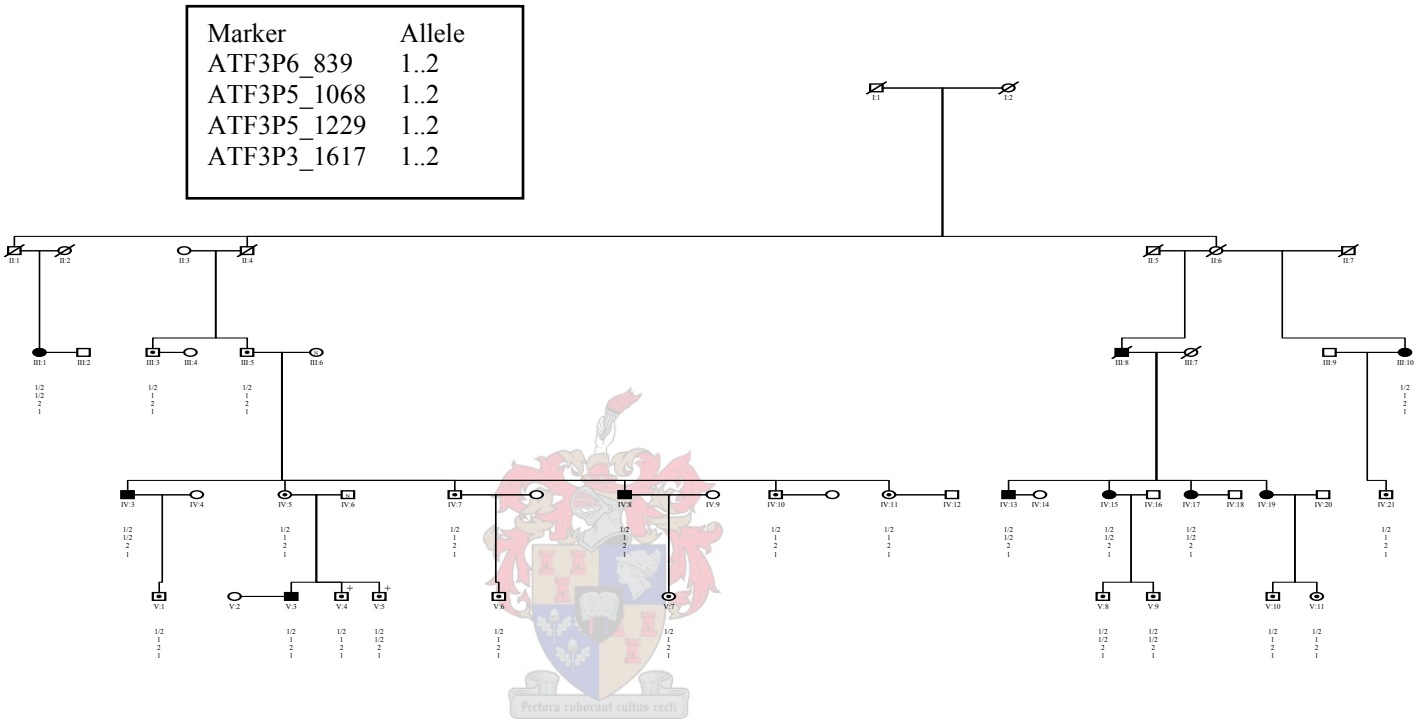


Figure 3.21 Pedigree 106, which was investigated for the segregation of single nucleotide changes identified in fragments 3, 5 and 6 of the ATF3 promoter region. *The family was reported in previous studies as carrying the* MYH7*-Arg403Trp mutation (Moolman et al., 1998; Moolman et al., 2000). Allele definition: ATF3P6_839- 1=T, 2=A; ATF3P5_1068- 1=G, 2=C; ATF3P5_1229- 1=C, 2=T; ATF3P3_1617- 1=A, G=2. 2= 2/2 and 1= 1/1. Symbol definition: Open symbols indicate individuals with unknown status; N indicates unaffected individuals; solid symbols indicate affected subjects; plus symbol indicates the* MYH7*-Arg403Trp mutation carriers, open symbol with a dot indicates the* MYH7*-Arg403Trp mutation carriers but not clinically affected; diagonal lines indicate deceased individuals. The order of the marker genotypes, from the telomere to the centromere, is indicated in the box at the top of the diagram.*

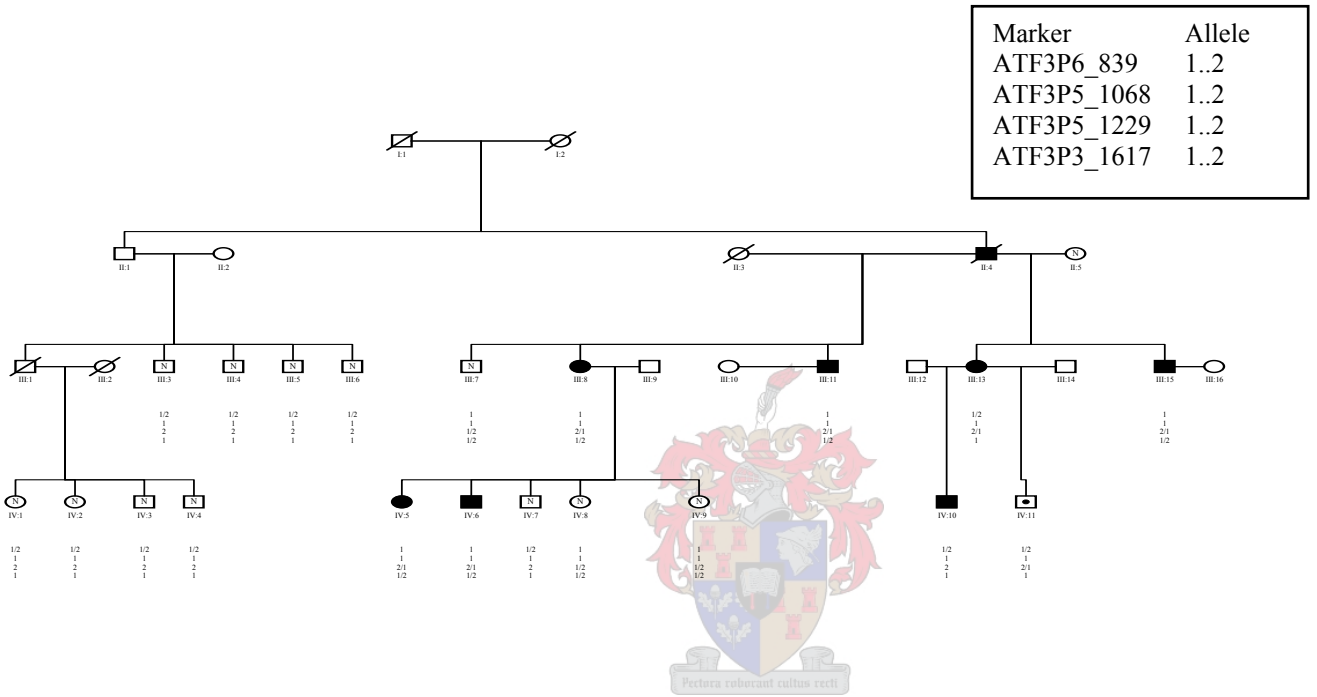


Figure 3.22 Pedigree 108, which was investigated for the segregation of single nucleotide changes identified in fragments 3, 5 and 6 of the* ATF3 *promoter region. *The family was reported in previous studies as carrying the* MYH7-Arg249Glu *mutation (Moolman et al., 1998; Moolman et al., 2000). Allele definition: ATF3P6_839- 1=T, 2=A; ATF3P5_1068- 1=G, 2=C; ATF3P5_1229- 1=C, 2=T; ATF3P3_1617- 1=A, G=2. 2= 2/2 and 1= 1/1. Symbol definition: Open symbols indicate individuals with unknown status; N indicates unaffected individuals; solid symbols indicate affected subjects; plus symbol indicates* MYH7-Arg249Glu *mutation carriers, open symbol with a dot indicates* MYH7-Arg249Glu *mutation carriers but not clinically affected; diagonal lines indicate deceased individuals. The order of the marker genotypes, from the telomere to the centromere, is indicated in the box at the top of the diagram.*

Fragment 8

Two different patterns were identified in individuals in the HCM panel, one in lane 1 and another in lane 2 (Fig. 3.23). Amplicons of individuals with these patterns were subsequently sequenced and an A→G transition substitution was identified at position -282 in both sequences obtained (Fig. 3.24). Genotypes of probands in the HCM panel and their family members could not be determined using ASREA because the variant did not affect a restriction enzyme recognition site.

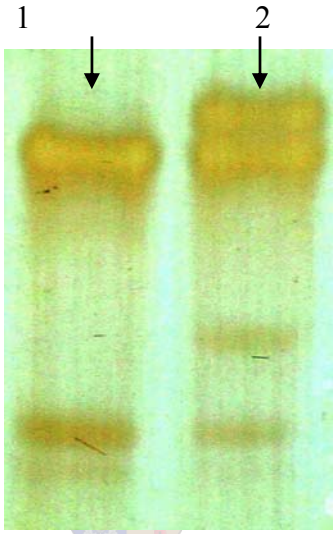


Figure 3.23 Representative PCR-SSCP analysis of* ATF3 *promoter region, fragment 8 in the HCM panel*. The arrows indicate variant patterns identified in HCM-affected individuals in the HCM panel.*

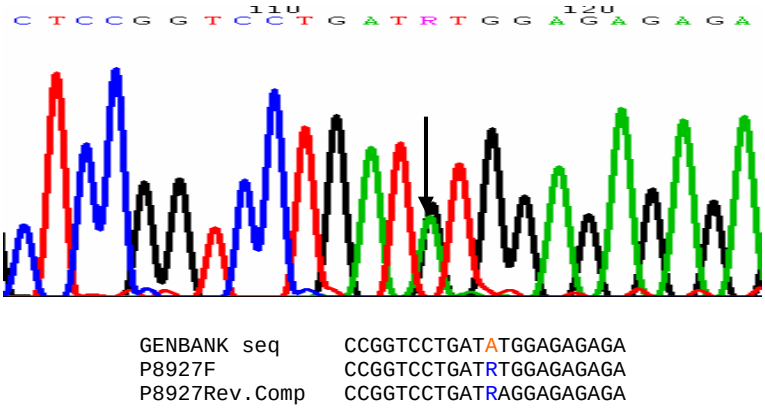


Figure 3.24Representative mutation analysis of* ATF3 *promoter region, fragment 8 in the HCM panel. *The -282A/G variant is indicated by an arrow. Below is the CLUSTAL alignment of the GENBANK ATF3 promoter sequence, and the forward and reverse complementary sequences obtained from the automated sequencing. R refers to the* A/G *heterozygote.*

3.4 Allele-specific restriction enzyme analysis

-1229C/T and -1068G/C variants

Two of sequence variations, -1068G/C or -1229C/T, identified in fragment 5 of the *ATF3* promoter region (Fig. 3.16) affected the restriction enzyme recognition site resulting in a gain of an *Nla*III site or a loss of an *Alw*26I site, respectively (Figs. 3.25 and 3.26) (Fig. 2.5 for a detailed restriction map). This digestion was therefore used to verify the genotypes of the analysed individuals.

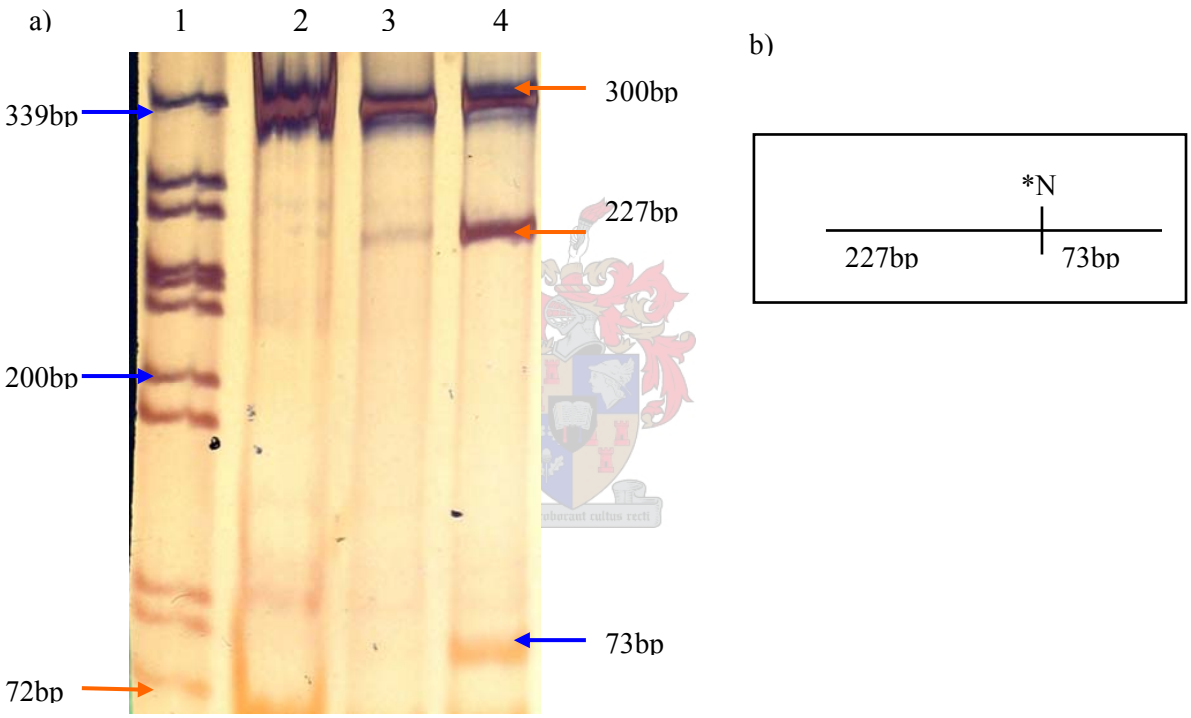


Figure 3.25 Genotyping of HCM-affected individuals at position -1068 of the* ATF3 *promoter for the G/C transition substitution. *The amplified promoter fragment was digested with Nla*III *and size-separated by electrophoresis on a 12% polyacrylamide gel.* The *C allele resulted in a gain of an Nla*III *recognition site as indicated by the presence of a 227-bp fragment. Lane 1 =* λPst*, lane 2 = undigested amplicon, lane 3 =* GG*, and lane 4 =* GC*.The faint band in lane 3 is caused by overflowing of the amplicon in lane 4. b) Schematic representation of the amplified fragment showing the position of the Nla*III *site. *= variable restriction enzyme site.*

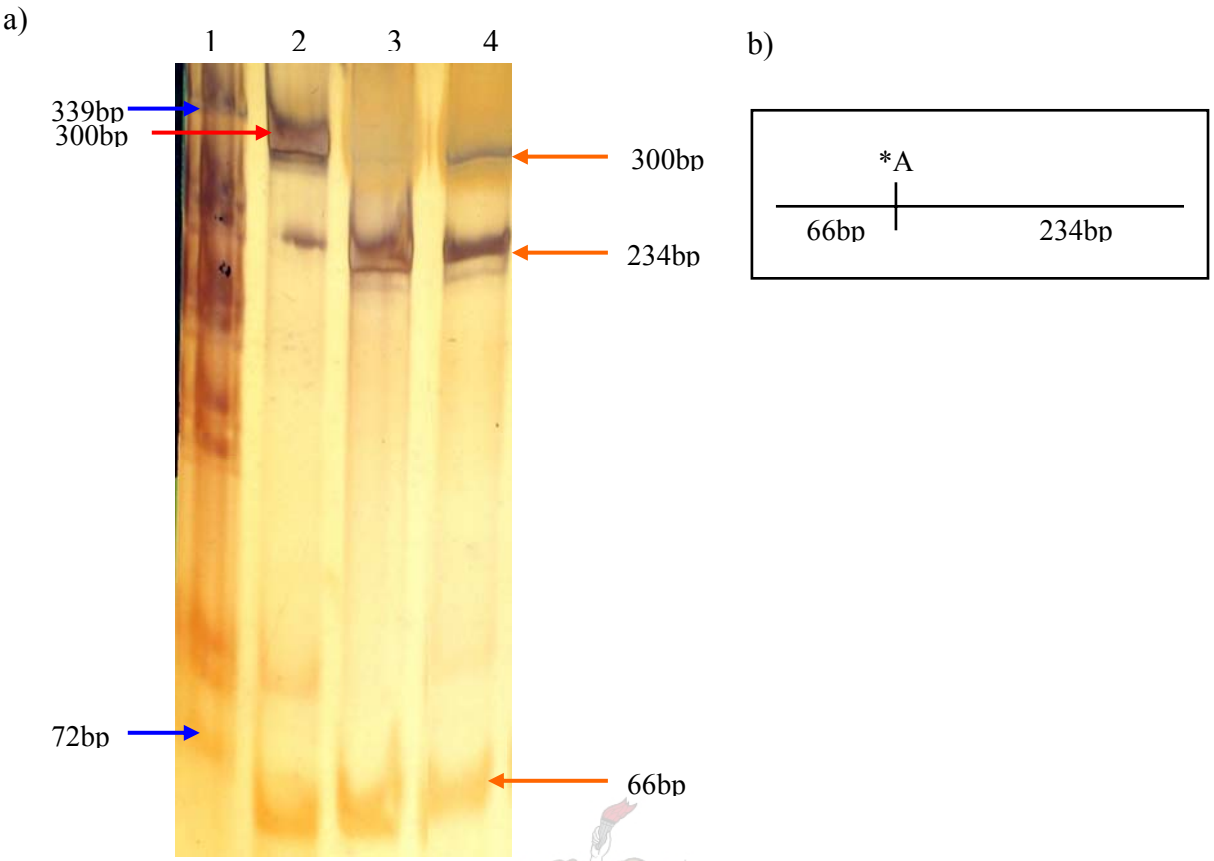


Figure 3.26 Genotyping of HCM-affected individuals at position -1229 of the ATF3 promoter for the C/T substitution. *a) The amplified promoter fragment was digested with an Alw*26I *and size-separated by electrophoresis on a 12% polyacrylamide gel.* The *T allele resulted in a loss of an Alw*26I *recognition site as indicated by the presence of the 300-, 234-, and 66-bp fragments in the* CT *genotype in lane 4. Lane 1 = λPst, lane 2 = undigested amplicon, lane 3 =* C/C*, and lanes 4 =* CT*. The second band in lane 2 is caused by overflowing of the amplicon in lane 3. b) Schematic representation of the amplified fragment showing the position of the Alw*26I *site. *= variable restriction enzyme site.*

3.5 Summary of mutation analysis of the ATF3 promoter region

In summary, from the genotype analysis of HCM-affected individuals in the HCM panel, the following data was obtained: the -1617*A/A* and -1617*A/G* genotypes were detected in fragment 3; the -1068*G/C*- -1229*C/C*, -1068*G/C*- -1229*C/T*, -1068*G/G*- -1229*C/C* and -1068*G/G*- -1229*C/T* genotypes in fragment 5 and the -839*T/T* and -839*T/A* genotypes in fragment 6.

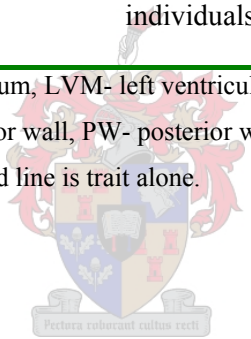
3.6 QTDT analysis

Transmission disequilibrium tests, such as QTDT, estimate whether preferential transmission of one allele at a polymorphic locus occurs between affected parent and affected children within a family (Abecasis *et al*., 2000). As such, these tests require that at least one of a set of parents in each nuclear family be heterozygous at the polymorphic locus. The QTDT program requires that at least 30 individuals in the given dataset be heterozygous before analysis commences. Although all *ATF3* promoter polymorphisms identified in the study were genotyped in the families of the probands (Figs. 3.20-3.22), only the -839T/A polymorphism was informative enough in these families to allow statistical analysis. The p-values obtained indicated that there is no significant association between this marker and the HCM-related traits (table 3.6). The analysis was corrected for age and gender as these covariates can affect the statistical values, and again no association was detected.

Table 3.6 *p values for TDT of the ATF3 -839*T/A *polymorphism and HCM-related quantitative traits.*

Marker/SNP	Quantitative trait	Number of individuals tested	P value (before correction)	P value (overall Bonferroni significant level)
-839T/A	LVM_age_gender	34 out of 76 individuals	0.2609	0.4537
	LVM		0.1423	0.2643
	IVS_age_gender	34 out of 76 individuals	0.5994	0.8395
	IVS		0.5010	0.7510
	LVPW_age_gender	34 out of 76 individuals	0.4464	0.6936
	LVPW		0.5284	0.7776
	LVED_age_gender	34 out of 76 individuals	0.4820	0.7316
	LVED		0.6678	0.8897
	PW+IVS_age_gender	34 out of 76 individuals	0.3357	0.5588
	PW+IVS		0.2978	0.5069

Abbreviations: IVS- interventricular septum, LVM- left ventricular mass, LVED- left ventricular end diastolic diameter, LVPW- left ventricular posterior wall, PW- posterior wall. The first line in each row is trait corrected for covariates age and gender. The second line is trait alone.



CHAPTER 4

DISCUSSION

INDEX

4.1. Identification of candidate genes within the PFHBII Locus

The present study sought to identify the molecular aetiology of PFHBII, a DCM-associated disorder complicated by cardiac conduction defects, which segregates in a South African Caucasian Afrikaner family (Brink and Torrington, 1977). To this end, plausible candidate genes were identified within the PFHBII locus and screened for a possible disease-causative mutation. In addition, these candidate genes were screened for HCM main locus and/or HCM modifying mutations based on the premise discussed in chapter 1 (section 1.7). Several strategies have been implemented to identify PFHBII candidate genes, of which two were sequence alignment analyses (BLAST) and data-mining of gene annotation databases (Ensembl and UCSC). Although no sequence homologies to the genes previously identified as causes of DCM, cardiac conduction disorders associated with arrhythmia or HCM were identified by BLAST homology searches at the PFHBII locus, eleven plausible PFHBII candidate genes were selected from this region using the Ensembl and UCSC gene identification tools (table 3.5).

Two genes, *PFKFB2* and *ATF3*, were screened for both PFHBII-causing and HCM main locus and/or modifying mutations using PCR-based SSCP and nucleotide sequence analyses. No PFHBII-causing mutation was identified upon analysis of subjects from the PFHBII-affected family in either *PFKFB2* or *ATF3*. Furthermore, recent database analysis indicated that *PFKFB2* and *ATF3* were mapped outside the PFHBII locus. Therefore, the two genes were excluded as cause of PFHBII by position, as well as by extensive mutation analysis of exons, exon/intron junctions and promoter sequences. Additionally, several sequence variants were identified in both candidate genes in the HCM panel: one variant was identified in the coding region of *PFKFB2*, and six variants were identified in the *ATF3* promoter region, however, they were not HCM main locus mutations. Insufficient data was generated in a pilot study to

draw conclusions about their role as HCM modifying mutations. Results obtained in the present study are discussed below.

4.1.1 Preparing a catalogue of genes previously identified as causes of HCM, DCM, and cardiac conduction disease

A catalogue of genes previously identified as cause of HCM, DCM, and conduction defects was prepared using published material and database resources. The catalogue consisted of the genes and associated diseases, the chromosomal locus of each gene, its genomic organisation and the accession number of the individual gene (refer to tables 3.1 and 3.2). Complete nucleotide and protein sequences of the genes in the catalogue are available in NCBI GenBank, Ensembl and UCSC, and can be retrieved using their accession numbers as identifiers. During the course of the study, the catalogue was continuously updated as new information about genes was obtained from published data and databases.

4.1.2 BLAST search

Sequences of the genes in the catalogue were used as queries to identify genes with homologous functional domains within the PFHBII locus. Although a number of gene identification strategies could be employed to identify candidate genes at a locus, it was opted to use BLAST homology searches in the present study. This approach was utilised because BLAST provides freely available, yet powerful, bioinformatics tools for performing functional analysis of a novel gene by comparing its sequence with that of genes that encode proteins with known structure and/or function (Altschul *et al*., 1990; Pearson, 1996). Genes previously identified as cause of DCM and HCM were used due to the fact that a primary form of cardiomyopathy is one of the clinical features of PFHBII (see section 1.3.3). Additionally, the identification of DCM-causing mutations in genes in which certain mutations had previously been reported as cause of HCM has resulted in speculation that there may be common

pathological pathways among these two forms of cardiomyopathy (Olson *et al*., 2001; Moolman-Smook *et al*., 2003). Therefore, we hypothesised that any gene within the PFHBII locus that is structurally and/or functionally related to the genes implicated as causing DCM or HCM is a plausible candidate for PFHBII mutation analysis.

In the present study, the BLAST analyses indicated that there were no genes within the PFHBII locus with homologous functional domains to those in the catalogue, or that they were missed as the program uses sequence similarities to identify related proteins. BLAST searches were performed after the release of the second draft sequence of the Human genome (June 2001) and again after the announcement of the completion of the HGP (14 April 2003), at the time when genes were still being annotated. Thus, it is possible that, when the searches were performed, a newly identified gene/protein signature was not yet entered into one of the HGP databases and, therefore, no reference to its unique sequence existed in the database.

The disadvantage of using the BLAST program to identify biologically related genes or proteins is its inability to identify distantly related genes or proteins (Brenner *et al*., 1998; Park *et al*., 1998). The limitation of the BLAST homology searches is also made apparent by its inability to identify the structural similarity between hemoglobin and myoglobin, which are considered to be biologically related although they have very low sequence similarity (Brenner *et al* ., 1998). It has been reported that protein classification is not only based on sequence identity, but also on structural details, folding of proteins and functional features (Park *et al*., 1998). Proteins with domains that have close evolutionary relationships (i.e., high sequence identity) are grouped into one family and those that have low sequence identity but whose structural details and functional features suggest an evolutionary origin, are grouped into superfamilies (Park *et al*., 1998). Although profile-based techniques such as PSI-BLAST (Altschul *et al*., 1997), Hidden-Markov models (Karplus *et al*., 1998) and structure-based

techniques such as ENVIRON (Bordo, 1993) were developed for identifying distantly related proteins, these tools are limited by the lack of structural and functional data of several newly annotated proteins in databases (Koppensteiner *et al*., 2000).

4.1.3 Ensembl and UCSC gene identification tools

These tools were used because, apart from providing information about a gene's chromosomal location, they also allow access to a lot of information related to the function of the gene's product and tissue expression profile.

4.1.3.1 Ensembl gene identification tools

At the start of the study (June 2001), two attractive candidate genes, *PFKFB2* and *ATF3*, were identified within the PFHBII locus and were therefore entered into the mutation screening. The rationale for selecting *PFKFB2* and *ATF3* as PFHBII and HCM plausible candidate genes is discussed below.

PFKFB2

Apart from apparently being a PFHBII plausible candidate gene by position at the start of the study, *PFKFB2* was chosen for its expression in cardiomyocytes and for its role in energy metabolism. *PFKFB2* encodes a PFK-2/FBPase-2, which catalyses the synthesis and degradation of fructose-2,6-bisphosphate (Fru-2,6-P_2). PFK-2/FBPase-2 regulates glycolysis through its substrate, Fru-2,6-P_2, that stimulates 6-phosphofructo-1-kinase (PFK-1) activity (Hasemann *et al*., 1996). PFK-1 is a key glycolytic enzyme, which catalyses the first reaction, phosphorylation of fructose-6-phosphate, that commits glucose to glycolysis (Depré *et al*., 1998). PFK-2/FBPase-2 has been reported to be a substrate for AMPK, an enzyme encoded by *PRKAG2*, the gene in which HCM-causative mutations have been identified (Blair *et al*., 2001). PFK-2 activation by AMPK is correlated with AMPK activation under anaerobic

conditions. The AMPK-mediated activation of PFK-2/FBPase-2 stimulates the flux through PFK-1 by increasing the concentration of Fru-2,6-P_2, which, in turn, stimulates ATP production through glycolysis. The involvement of protein phosphorylation in the Pasteur effect is superimposed on the direct stimulation of PFK-1 by an increase in intracellular AMP:ATP ratio (Marsin *et al*., 2000). This indicates that PFK-1 stimulates glycolysis through two mechanisms: a direct allosteric stimulation by an increase in the AMP:ATP ratio and an indirect mechanism involving the phosphorylation of PFK-2/FBPase-2 by AMPK.

AMPK has also been reported to play a role in the regulation of ATP production. The enzyme acts as a sensor of the energy state through its activation, like PFK-1, by an increase in the intracellular AMP:ATP ratio. AMPK phosphorylates and inactivates enzymes in energy-consuming biosynthesis pathways, such as synthesis of fatty acids, cholesterol, glycogen and phosphocreatine, to conserve ATP. It has been suggested that alterations in cellular energy homeostasis may be the underlying pathophysiolocal mechansism in DCM and HCM (Ashrafian *et al*., 2003). Therefore, because of its role in maintaining cellular energy homeostasis, *PFKFB2* was selected as a plausible PFHBII candidate gene.

ATF3

Apart from initially being a PFHBII plausible candidate gene by its apparent position within the PFHBII locus, *ATF3* was chosen for its expression in cardiomyocytes and for its role in eliciting hypertrophic response in cardiac tissue of transgenic mice models (Okamoto *et al.*, 2001). *ATF3* encodes an activating transcription factor 3 (ATF3), a member of the CREB/ATF family of basic region-leucine zipper transcription factors (Brindle and Montminy, 1992). ATF3 mRNA levels are increased in the heart after myocardial ischemia-reperfusion, in the kidneys after renal ischemia-reperfusion, in the skin after wounding, in the brain after seizure and in the liver after chemical toxicity and partial hepatectomy (Chen *et al*., 1996; Yin *et al*.,

1997). *In vitro* studies performed with cultured cells also indicated that *ATF3* is induced by cytokines (Drysdale *et al*., 1996; Farber, 1992), genotoxic agents and agents that are known to induce cell death (Amundson *et al*., 1999; Liang *et al*., 1996). The induction of *ATF3* by myocardial ischemia-reperfusion was also demonstrated in *in vivo* studies performed with rat models, in which the animals were treated with coronary artery ligation coupled with reperfusion, to mimic myocardial ischemia-reperfusion (Okamoto *et al.*, 2001). Myocardial ischemia-reperfusion in the rat model increased the mRNA levels of *ATF3*. Okamoto and co-workers also investigated the effect of elevated *ATF3* mRNA levels in transgenic mice models. Elevated *ATF3* mRNA levels resulted in bi-atrial enlargement and an increase in heart-body weight, which was due to cardiac hypertrophy. In addition to hypertrophy, the transgenic mice showed atrial and right ventricular dilation. Myocyte disarray, degenerating mitochondria, abnormal Z-lines, vacuoles, granules and degenerating intercalated disks were also noted. The transgenic mice also demonstrated altered AV node conduction. These studies indicate that expression of *ATF3* in the heart results in conduction abnormalities and cardiomyopathies, suggesting that *ATF3* induction by stress signals may play a role in the pathogenesis of stress-associated cardiac diseases.

Another member of the CREB/ATF family, the cAMP-responsive element-binding protein (CREB), has also been implicated in cardiac diseases (Fentzke *et al*., 1998; Leiden, 1997). Transgenic mice expressing a dominant-negative form of CREB have been shown to develop four-chambered DCM. These findings indicate that CREB is important for normal functioning of the heart, while the studies performed on *ATF3* suggest that this gene may play a role in cardiac stress responses, as it is not detected in the heart under nonstressed conditions. *ATF3* has also been reported to be a downstream target for the c-Jun N-terminal and p38 stress kinases, which are implicated to play a role in promoting cardiac pathology, including myocyte hypertrophy and apoptosis (Yue *et al*., 2000; Turner *et al*., 1998).

Although *PFKFB2* and *ATF3* were initially chosen as PFHBII plausible candidate genes, they were later mapped outside the PFHBII locus due to errors in the order of markers on physical maps, which can be caused by problems with assembly of bacterial artificial chromosome (BAC) clones. Such problems were noted not only in chromosome 1, but in other chromosomes as well (DeWan *et al*., 2002). For example, in the draft human genome sequence released on October 2000, a clone (AC037449.2) derived from chromosome 10 that contains several markers (*D10S1225*, *D10S2278*, *SHGC-130658*, *stSG29297* and *stSG43409*) was misassembled onto chromosome 14 (DeWan *et al*., 2002). BAC libraries were used in human genome sequencing as a source of substrate for shotgun sequencing project and restriction fingerprints for building overlapping clone sets (International Human Genome Sequencing Consortium, 2001). BACs also provided information for mapping sequence contigs to localised genomic regions by using a direct genomic shotgun sequencing approach. However, the disadvantage of using BACs is the occurrence of rearrangement in approximately 10% of the clones, which can result in errors in the order of markers (Osoegawa *et al*., 2001). Further problems in assembly can be due to the location of markers near gaps in physical-sequence data and to errors in *in silico* mapping. Filling of gaps in the human genome sequence and gene annotation changes have been shown to occur frequently throughout the progress of the HGP (DeWan *et al*., 2002).

4.1.3.2 UCSC gene identification tools

UCSC gene prediction tools were also used to identify other strong candidate genes harboured at the PFHBII locus after the announcement of the completion of the HGP. These genes were chosen as plausible PFHBII candidates based on predicted functions of their products. The proteins encoded by these active genes are predicted to have functions that include regulation of apoptosis, nucleic acid (DNA) binding, ATP-dependent transportation of molecules in the

mitochondrion (mitochondrial carrier proteins), plasma membrane calcium transportation, transcription regulation, signal transduction, and phospholipid metabolism. The rationale for their selection for future mutation screening is discussed below.

Regulation of apoptosis- G0S2

The putative lymphocyte G0/G1 switch gene (*G0S2*) is predicted to play a role in regulation of apoptosis. Apoptosis is one of the late changes in end-stage heart failure, which include changes in gene expression profile, contractile dysfunction and interstitial remodeling in response to persisting cardiac stress (Petrich and Wang, 2004). The origin of heart failure ranges from hypertension and coronary artery diseases to genetic defects, with cardiac hypertrophy developing as a compensatory response that improves the heart's contractility and function (reviewed by Frey and Olson, 2003). The role of apoptosis in heart failure has been supported by the upregulation of genes involved in apoptosis (e.g., voltage-dependent anion channel 1 and 2, prostaglandin synthase D, and kallikrein 11), which was detected in a gene expression profile study performed with sample biopsies of DCM-affected individuals (Grzeskowiak *et al.,* 2003).

Mitochondrial carrier proteins- FLJ25078

The *FLJ25078* gene encodes a protein that is functionally related to the Ca^{2+}- and ATP-binding mitochondrial carrier proteins. Mitochondria carry out various processes that are important for cellular development and survival (reviewed by Wallace, 2001). These processes include energy generation (citric acid cycle and fatty acid β-oxidation), amino acid synthesis and degradation (urea cycle), the synthesis of iron–sulphur clusters and haem, and the generation of heat by dissipation of the proton gradient. To achieve this, the mitochondria must transport solutes between the mitochondrial matrix and the cytoplasm. Many of the transport

steps are carried out over the inner mitochondrial membrane by the mitochondrial carrier proteins.

Transcription regulation and signal transduction- RAMP zinc finger protein 198, IRF6 and TRAF5

The interferon receptor factor 6 (IRF6) and RA-regulated nuclear matrix-associated protein (RAMP)-zinc finger proteins are predicted to play a role in regulation of transcription (Mamane *et al*., 1999). The tumor necrosis factor receptor-associated 5 (TRAF5) is predicted to play a role in signal transduction (Song *et al*., 1997). Signal transduction and re-induction of gene expression have been reported to be among of the changes that occur due to an increase in biochemical stress in cardiomyocytes (reviewed by Frey and Olson, 2003). Cardiac hypertrophy develops in response to this increase in biochemical stress, which can be caused by extrinsic factors such as hypertension or valvular heart disease, or intrinsic factors such as gene defects. Cardiac hypertrophy is usually preceded and accompanied by re-induction of transcription, the downsteam target of Ca^{2+}-dependent signal transduction pathways that are activated through the G protein-coupled receptors (Frey and Olson, 2003). The role of transcription factors in cardiac disorders is also supported by the fact that *PRKAG2* has been implicated as causing HCM that is complicated with conduction defects (section 1.4.3 and 1.5.2). Molecular characterisation of AMPK revealed the existence of mammalian AMPK homology to the yeast SNF-1 and plant SnRK transcription factors, although its gene expression regulation in mammals has not been investigated (Carlson, 1999; Halford and Hardie, 1998; Alderson *et al*., 1991).

Phospholipid metabolism- KIAA0205

The protein encoded by *KIAA0205* is predicted to play a functional role in phospholopid metabolism. The role of proteins involved in phospholipid metabolism in cardiac function is

supported by the fact that energy generation in the heart is dependent on ATP production by mitochondrial oxidation of long-chain fatty acids (Randle *et al*., 1963). In hypertrophic cardiomyocytes, fatty acid oxidation is suppressed and glucose utilisation favored. Genes involved in fatty acid oxidation are regulated by a family of transcription factors, the peroxisome proliferator-activated receptors (PPARs) (Barger and Kelly, 2000). Functional studies performed on transgenic mice with a mutant PPAR gene demonstrated embryonic lethality due to myocardial defects (Barak *et al*., 1999). Genes involved in cholesterol synthesis (e.g., hydroxymethylglutaryl coenzyme A reductase) have also been implicated in hypertrophic signalling (Laufs *et al*., 2002). These findings suggest that genes and proteins involved in lipid metabolism may play a role in cardiac abnormalities.

Gene expression profiling in DCM and HCM

Gene expression studies performed on cardiac tissue biopsies of DCM- and HCM-affected individuals showed changes in expression levels of a variety of genes (Lim *et al*., 2001, Grzeskowiak *et al*., 2003). Although *PFKFB2* and *ATF3* were not specifically identified in these studies, the classes of genes in which changed expression profiles were reported included those encoding contractile sarcomeric and cytoskeletal proteins, ion channels, intracellular signal transducers, proteins maintaining the redox state of the myocardium, along with transcriptional and translational machinery (Lim *et al*., 2001, Grzeskowiak *et al*., 2003).

The expression profile studies provided a resource to investigate if any of the genes identified, or those with analogous functions, map to the PFHBII locus and, therefore, are candidate genes to screen in future mutation analysis.

4.2 Mutation analysis of candidate genes in the PFHBII family

4.2.1 Search for PFHBII-causing mutations in the PFHBII family

4.2.1.1 PFKFB2

PFKFB2 was screened for PFHBII-causing mutations because, at the time of the present study, it was localised within the PFHBII locus as indicated in figure 3.1. No PFHBII-causing mutation was identified in this gene. The different electrophoretic band patterns identified in exon 8 did not represent sequence variants, as there were no nucleotide changes detected in sequenced amplicons. The reason for this is unknown since no paper has been published that explains the presence of mobility shifts that are not the result of sequence changes in the SSCP gels. Possible explanations may be that the extra bands detected were either PCR artefacts, or alternative stable conformers of the wild type sequence that were formed due to the exposure of samples to subtle variations in salt concentrations (de Lange, Phd thesis).

Mutations have been reported in the promoter regions of several genes that may either increase or decrease the expression level of a gene (Papassotiropoulos *et al*., 2001; Pola *et al*., 2002; Ross *et al*., 2004). A change in the expression level of a gene may influence the development of a particular disease. For example, polymorphisms in the promoter regions of the IL6 and TNFα genes were reported to affect gene expression levels, which has been observed in individuals with Alzheimer's disease (Papassotiropoulos *et al*., 2001; Pola *et al*., 2002). Therefore, a 500bp-*PFKFB2* promoter region was also screened for the PFHBII-causing mutation but no disease-causing mutation was identified. The reason for analysing only this region of the promoter is that five transcription-binding sites (three *Sp1*- and two *HNF1*-binding sites) were identified only in this part of the *PFKFB2* promoter through alignment with the *R. norvegicus pfkfb2* promoter region (Fig 3.4) (Heine-Suner *et al*., 1998), while none has been reported in the remaining 5' untranslated region of the gene. However, the role of these transcription-binding sites in the regulation of *PFKFB2* is unknown since their role in

promoter activity has not been studied. *PFKFB2* was excluded as a plausible candidate for PFHBII-causing mutation because it was later mapped outside the PFHBII locus.

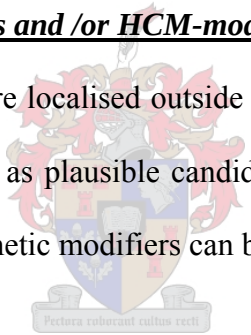
4.2.1.2 ATF3

No PFHBII-causing mutation was identified in the three exons of *ATF3* that encode its functional domains. The *ATF3* promoter region was also analysed for PFHBII-causing mutations in the PFHBII family and no PFHBII-causing mutation was identified. *ATF3* was later localised outside the PFHBII region, and therefore, also excluded as a strong candidate gene for PFHBII-causing mutations.

4.3 Mutation analysis of candidate genes in the HCM panel

4.3.1 Search for HCM main locus and /or HCM-modifying mutations

Although *PFKFB2* and *ATF3* were localised outside the PFHBII locus during the later stages of the project, they still qualified as plausible candidate genes for HCM causal and/or HCM modifying mutations, as HCM genetic modifiers can be located in any chromosomal region.



4.3.1.1 PFKFB2

Analysis of *PFKFB2* for possible HCM main locus and/or modifying mutations revealed a 11535G/A transition substitution in exon 5 of one HCM-affected individual (Figs 3.9 and 3.10). However, this variant affected neither an amino acid residue nor a splice junction, and, therefore, is most unlikely to have any effect on HCM phenotype. A 500bp-promoter region of *PFKFB2* was also screened for HCM main locus and/or modifying mutations but no sequence variants were identified (Fig. 3.11). However, the remaining part of the promoter region still requires mutation analysis to make conclusions regarding the potential role of *PFKFB2* in modifying the phenotypic expression of HCM.

4.3.1.2 ATF3

The *ATF3* coding region was analysed for HCM-causing and/or modifying mutations and no sequence variants were identified. The promoter region of the gene was also screened for mutations that may cause HCM or modulate its phenotypic expression. Mobility shifts were identified in three fragments of the *ATF3* promoter region (Figs. 3.13, 3.15, 3.18 and 3.23). Automated sequencing of three amplicons with different electrophoretic band patterns revealed a sequence variant (-1617*A/G* transition substitution) and a C nucleotide at position -1688 (not a T as detected in the published sequence) in fragment 3 of the *ATF3* promoter region (Fig. 3.14). However, whether the C or T nucleotide was present in other individuals that were analysed by PCR-SSCP method is unknown, as only three amplicons were sequenced. No additional sequence variant was identified in the third amplicon. Genotyping of affected individuals in the HCM panel and their family members revealed two genotypes in fragment 3 at position -1617: *AA* homozygotes and *AG* heterozygotes. The genotypes identified in fragment 3 could not be confirmed by ASREA, as these variants did not affect restriction enzyme recognition sites.

Three electrophoretic band patterns were detected in fragment 5 of the *ATF3* promoter region (Fig 3.15). Automated sequencing of the amplicons with different patterns revealed two sequence variants (Figs. 3.16 and 3.17). Amplicons of fragment 5 were re-analysed by ASREA to confirm the presence of sequence variants and to determine the genotype of other affected individuals in the HCM panel and their family members, as genotyping could not be performed by PCR-SSCP due to difficulties in interpreting different electrophoretic banding patterns. Several genotypes were identified in fragment 5 at positions -1068 and -1229, respectively, among the HCM probands and their family members: *G/C-C/C*, *G/C-C/T*, *G/G-C/C* and *G/G-C/T* (Figs. 3.24 and 3.25).

Two other variants were identified in fragments 6 and 8: -839T/A in fragment 6 (3.19) and -282A/G in fragment 8 (Fig. 3.23). Genotyping of individuals in the HCM panel and their families performed using PCR-SSCP analysis revealed two genotypes at position -839, the *TT* and *TA* genotypes. The genotypes could not be confirmed by ASREA, as these variants did not affect restriction enzyme recognition sites. Although two amplicons showed different electrophoretic mobility patterns on an SSCP gel, sequencing revealed that both were -282A/G heterozygotes. It was also not possible to confirm the genotypes of other individuals in the HCM panel and families using ASREA as the variant did not affect a restriction enzyme recognition site.

The data showed that the presence of the *ATF3* promoter region -282A/G variant could not be consistently detected as a mobility shift on SSCP gels. The sequence variant was only reliable detected by automated sequencing. This confirms the findings reported in previous studies that the PCR-SSCP analysis can detect approximately 80% of sequence variants (Sheffeld *et al*., 1993). Although in the present study, the sensitivity of the method was increased to approximately 98% by using different conditions (8% and 10% mild denaturing conditions), it has been shown that a small percentage was still not accounted for in experiments using known mutations under these conditions (de Lange, PhD thesis).

The sequence variants identified in the *ATF3* promoter region were found in both affected and unaffected individuals in probands HCM panel and their family members, therefore, they were not HCM main locus mutations. Futhermore, the variants did not affect any of the transcription-binding sites within the region, although the -1617A/G variant detected in fragment 3 is located one nucleotide away from the Myc/Max-binding site, one of the sites including E2F, implicated in cell cycle regulation (Liang *et al*., 1996). It is not clear, however, whether the Myc/Max site is functionally important for the *ATF3* promoter activity. The *ATF3*

promoter region has been reported to contain various transcription factor-binding sites, belonging to two groups: the inducible sites and sites implicated in cell cycle regulation (Liang *et al*., 1996). Inducible sites include ATF/CREB, AP1 and NF-$_k$B sites, which were demonstrated to be induced by signals such as cAMP, calcium influx, UV irradiation and cytokines (Meyer and Habener, 1993; Roesler *et al*., 1988).

In the present study, no functional studies were performed to determine the effect of different genotypes identified on *ATF3* promoter activity and gene expression level. Previous studies performed with *ATF3* in transgenic mice models demonstrated cardiac abnormalities caused by elevated *ATF3* mRNA levels (Okamoto *et al.*, 2001). However, QTDT analysis was performed in an attempt to determine the association between the identified genotypes and HCM quantitative phenotypes that are used in the disease diagnosis.

4.4 QTDT analysis

QTDT is a family based test that is used to determine linkage between alleles and phenotypes and the increased transmission of one allele at a polymorphic locus between affected parent and affected children within a family (Abecasis *et al*., 2000). The test was used in the present study as a pilot investigation of the possible association between the genotypes identified in the *ATF3* promoter region and four HCM-related quantitative traits (LVM, IVS, LVPW and LVED) that were measured by echocardiographic analysis as HCM diagnostic criteria (Maron *et al*., 1978). However, statistical analysis could only be performed for one variation, -839T/A, as the QTDT program requires sufficient numbers (at least 30) of heterozygous individuals before commencing analysis. Only the -839T/A gene variant was sufficiently informative in the HCM families to perform the analysis and no significant p-values were obtained as an indication of the presence of association between the alleles and HCM traits (table 3.6).

Additional clinical data for other traits (e.g., body mass index and blood pressure), that are equally important, were not available to pursue the analysis further.

At this stage, it is difficult to draw any conclusions about whether or not *ATF3* is a modifier of HCM phenotype because of insufficient QTDT data, and the fact that functional studies were not performed to determine the effect of the genotypes identified in the promoter region on the expression level of the gene.

4.5 Future studies

The two genes, *PFKFB2* and *ATF3*, were excluded as the PFHBII-causative gene because the disease–causative mutation was not identified upon extensive analyses of these genes, in addition to the fact that they were later mapped outside the PFHBII locus. Other candidate genes still need to be screened for both PFHBII-causing and HCM-causing and/or modifying mutations. Mutation screening of these genes and determination of their possible role in cardiac tissue will give insight on mechanisms underlying PFHBII and HCM. A molecular cause of PFHBII still remains to be identified as no causal mutation was identified in the present studies and none has been reported by other investigators yet (Fernandez, PhD thesis 2004). The genetic causes of HCM are not fully understood as the disease-causative genes identified to date only account for approximately 60% of FHC cases (Watkins, 2003). There is still a mystery surrounding the genetic factors that may contribute to the development and phenotypic expression of HCM. Although several genes have been reported that may play a role in modulating the disease phenotype, SNP association studies performed with these genes have given conflicting data that which may be due to linkage among polymorphisms within several genes.

Identification of the PFHBII-causative gene will allow a DNA-based diagnosis that will surpass the current use of genetic markers, which is limited to a family setting. It will also help in elucidation of the underlying molecular mechanisms that cause PFHBII. Furthermore, identification of both HCM-modifying and HCM-causing genes will give insight into the phenotypic variability noted among South African HCM-affected individuals and into the molecular cause of the disease among individuals with HCM-like clinical features. This may lead to the development of more directed therapies that may have application to both familial and sporadic forms of cardiac abnormalities in the general population.

APPENDIX I

URLS OF THE WEBSITES USED IN THE STUDY

http://www.cotc.edu/vstone/webdocek.html

http://www.cotc.edu/vstone/webdocek.html

http://www.embl.de/~wendt/ troponin.html

http://www.ensembl.org/Homo_sapiens/stats/status.html

http://www.ncbi.nlm.nih.gov

http://www.ncbi.nlm.nih.gov/BLAST

http://www.ensembl.org/

http://www.ensembl.org/Multi/blastview?species=Homo_sapiens

http://www.hopkinsmedicine.org/cardiology/ heart/photographs.html

http://cardiogenomics.med.harvard.edu/project-detail?project_id=230

http://www.sph.umich.edu/csg/abecasis/QTDT/

http://www.ebi.ac.uk/emboss/align/index.html

APPENDIX II

BUFFERS, MARKERS AND SOLUTIONS

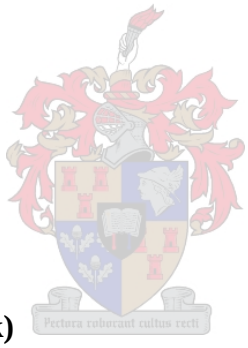
1 BUFFERS

1.1. CELL LYSIS BUFFER

Sucrose	0.32M
Triton-X-100	1%
$MgCl_2$	5mM
Tris-HCl	10mM
H_2O	1l

1.2. DNA EXTRACTION BUFFER

NaCl	0.1M
Tris-HCl	0.01M
EDTA (pH8)	0.025M
SDS	0.5%
Proteinase K	0.1mg/ml



1.3. TBE-BUFFER (10x stock)

Tris-HCl	0.89M
Boric Acid	0.89M
Na_2EDTA (pH8)	20mM

1.4. TE-BUFFER (10x stock)

TrisOH	0.1M
EDTA(pH8)	0.01M
H_2O	150ml

1.5. *Taq* POLYMERASE BUFFER (Bioline)

Ammonium sulphate	160mM
Tris-HCl (pH 8.8)	670mM
Tween-20	0.1%

2. SOLUTIONS

2.1. SOLUTIONS FOR POLYACRYLAMIDE GELS

2.1.1. Acrylamide-Bis (ready-to-use) solution 40% (37:5:1)

2.1.2. 10% AMMONIUMPERSULPHATE (APS)

APS	2g
H_2O	20ml

2.2. SOLUTIONS FOR SILVER STAINING

2.2.1. 0.1% $AgNO_3$ (Solution B)

$AgNO_3$	1g
H_2O	1L

2.2.2. DEVELOPING SOLUTION (Solution C)

NaOH	15g
$NaBH_4$	0.1g
Formaldehyde	4ml
H_2O	1L

3. GELS

3.1. 12% POLYACRYLAMIDE GEL

40% Acrylamide-Bis (ready-to-use) solution	3ml
10xTBE	1ml
Distilled H_2O	6ml
APS	80µl
TEMED	30µl

3.2. 10% MILDLY DENATURING POLYACRYLAMIDE GEL WITH 5% GLYCEROL

40% Acrylamide-Bis (ready-to-use) solution	40.5ml
10xTBE	8ml
Glycerol	8ml
Urea	24g
Distilled H_2O	84ml
APS	1600µl

TEMED	160µl

3.3. 8% MILDLY DENATURING POLYACRYLAMIDE GEL WITH 5% GLYCEROL

40% Acrylamide-Bis (ready-to-use) solution	32.4ml
10xTBE	8ml
Glycerol	8ml
Urea	24g
Distilled H_2O	91.8ml
APS	1600µl
TEMED	160µl

4. LOADING DYES

4.1. BROMOPHENOL BLUE

Bromophenol blue	0.2%(w/v)
Glycerol	50%
Tris (pH8)	10mM

4.2. SSCP LOADING DYE

Formamide	95%
EDTA	20mM
Bromophenol blue	0.01%
Xylene cyanol	0.05g
NaOH	10mM

5. MOLECULAR WEIGHT MARKER (LAMBDA *Pst*I)

Bacteriophage Lambda DNA (250µg)	100µl
Buffer M (Boerhinger Mannhein)	15µl
*Pst*I (Boerhinger Mannheim)	11µl
H_2O	32µl

Incubate at 37°C for 2 hours followed by heat inactivation at 65°C for 5 minutes. Load 2µl onto polyacrylamide gels.

APPENDIX III

LIST OF SUPPLIERS

Acrylamide-Bis (ready-to-use) solution	Promega
$AgNO_3$	Merck
Alw 26I	Sigma
Ammonium persulphate	Merck
Boric acid	Merck
Bromophenol blue	Merck
dGTP	Boerhinger Mannheim
dCTP	Boerhinger Mannheim
dTTP	Boerhinger Mannheim
dATP	Boerhinger Mannheim
*EcoR*V	Promega
EDTA	Boerhinger Mannheim
Ethanol	Boerhinger Mannheim
Formamide	Merck
Formaldeyde	Merck
Gelbond	Merck
GFX™ PCR DNA and Gel Band Purification Kit	Amersham Bioscience
Glycerol	FMC
K-acetate	Sigma
KCl	Merck
Lambda DNA	Promega
Mineral oil	BDH Chemicals
NaAc	Merck
NaCl	BDH Chemicals
NaOH	Sigma
*Nla*III	New England Biolabs
Phenol	Merck
PBS	Sigma
SDS	Sigma
Taq polymerase	Bioline

TEMED	Sigma
Tris	Merck
Tris-OH	Merck
Tris-HCl	Merck
Urea	BDH Chemicals

REFERENCE

Aaronson RP, Blobel G. (1975) Isolation of nuclear pore complexes in association with a lamina. *Proc. Natl Acad. Sci. USA* **72:**1007–1011.

Abecasis GR, Cardon LR, Cookson WO. (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet*. **66**:279-292.

Abelman WH, and Lorrell BH (1989) The challenge of cardiomyopathy. *J. Am. Coll. Cardiol.* **13**: 1219-1239.

Adams, T D; Yanowitz, F G; Fisher, A G; Ridges, J D; Nelson, A G; Hagan, A D; Williams, R R; Hunt, S C. (1995) Heritability of cardiac size: an echocardiographic and electrocardiographic study of monozygotic and dizygotic twins. *Circulation* **71**:39-44.

Akai J, Makita N, Sakurada H, Shirai N, Ueda K, Kitabatake A, Nakazawa K, Kimura A, Hiraoka M. (2000) A novel SCN5A mutation associated with idiopathic ventricular fibrillation without typical ECG findings of Brugada syndrome. FEBS Lett. **479**:29-34.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994). The cytoskeleton. Robertson A. and Adams R. (3rd ed.), Molecular Biology of the Cell (847-857). New York: Garland Publishers Inc.

Alderson A, Sabelli PA, Dickinson JR, Cole D, Richardson M, Kreis M, Shewry PR, Halford NG. (1991) Complementation of snf1, a mutation affecting global regulation of carbon metabolism in yeast, by a plant protein kinase cDNA. *Proc Natl Acad Sci USA*. **88**:8602-8605.

Alings M, Wilde A. (1999) "Brugada" syndrome: clinical data and suggested pathophysiological mechanism. *Circulation* 99:666–673.

Allison DB, Heo M, Kaplan N, Martin ER. (1999) Sibling-based tests of linkage and association for quantitative traits. *Am J Hum Genet*. **64**:1754-1763.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. **25**:3389-3402.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990) Basic local alignment search tool. *J Mol Biol*. **215**:403-410.

Amundson SA, Bittner M, Chen Y, Trent J, Meltzer P, Fornace AJ Jr. (1999) Fluorescent cDNA microarray hybridization reveals complexity and heterogeneity of cellular genotoxic stress responses. *Oncogene*. **18**:3666-3672.

Anan R, Greve G (1994) Prognostic implication of novel β cardiac myosin heavy chain gene mutations that cause familial hypertrophic cardiomyopathy. *J Clin Invest* **93**:280-285.

Anan R, Shono H, Kisanuki A, Arima S, Nakao S, Tanaka H. (1998) Patients with familial hypertrophic cardiomyopathy caused by a Phe110Ile missense mutation in the cardiac troponin T gene have variable cardiac morphologies and a favorable prognosis. *Circulation.* **98**:391-397.

Arad M, Benson DW, Perez-Atayde, *et al*. (2002) Constitutively active AMP kinase mutations cause glycogen storage disease mimicking hypertrophic cardiomyopathy. *J Clin Invest*. **109**:357-362.

Arber S, Halder G, Caroni P (1994) Muscle LIM protein, a novel essential regulator of myogenesis, promotes myogenic differentiation. *Cell* **79**:221-231.

Arbustini E, Diegoli M, Morbini P, Dal Bello B, Banchieri N, Pilotto A, Grasso M, Narula J, Gavazzi A, Vigano M, Tavazzi L (2000) Prevalence and characteristics of dystrophin defects in adult male patients with dilated cardiomyopathy. *J Am Coll Cardiol* **35**:1760-1768.

Arbustini E, Pilotto A, Repetto A, Grasso M, Negri A, Diegoli M, Campana C, Scelsi L, Baldini E, Gavazzi A, Tavazzi L. (2002) Autosomal dominant dilated cardiomyopathy with atrioventricular block: a lamin A/C defect-related disease. *Am Coll Cardiol*. **39**:981-90.

Ashrafian H, Redwood C, Blair E, Watkins H. (2003) Hypertrophic cardiomyopathy: a paradigm from myocardial energy depletion. *Trends in Genetics* **19**:263-268.

Bang ML, Centner T, Fornoff F, *et al*. (2001) The complete gene sequence of titin, expression of an unusual ≈700-kDa titin isoform and its interaction with obscurin identify a novel Z-line to I-band linking system. *Circ Res***. 89**:1065–1072.

Barak Y, Nelson MC, Ong ES, Jones YZ, Ruiz-Lozano P, *et al*. (1999) PPAR gamma is required for placental, cardiac, and adipose tissue development. *Mol. Cell* **4**:585-595.

Barger PM, Kelly DP. (2000) PPAR signaling in the control of cardiac energy metabolism. *Trends Cardiovasc. Med.* **10**:238-245.

Baroudi G, Pouliot V, Denjoy I, Guicheney P, Shrier A, Chahine M. (2001) Novel mechanism for Brugada syndrome: defective surface localization of an SCN5A mutant (R1432G). *Circ Res* **88**:E78–E83.

Barton PJ, Cullen ME, Townsend PJ, Brand NJ, Mullen AJ, Norman DA, Bhavsar PK, Yacoub MH. (1999) Close physical linkage of human troponin genes: organization, sequence, and expression of the locus encoding cardiac troponin I and slow skeletal troponin T. *Genomics*. **57**:102-109.

Baxevanis AD. (2001) The Molecular Biology Database Collection: an updated compilation of biological database resources. *Nucleic Acids Res*. **29:**1-10.

Beckerle MC. (1986) Identification of a new protein localized at sites of cell-substrate adhesion. *J Cell Biol*. **103**:1679-1687**.**

Belkin A, Ornatsky O, Kabakov A *et al*. (1988) Diversity of vinculin/metavinculin in human tissues and cultivated cells: expression of muscle specific variants of vinculin in human aorta smooth muscle cells. *J Biol Chem*. **263**:6631-6635.

Bennet BP, Yazawa K, Makita N, George AL, Jr. (1995) Molecular mechanism for an inherited cardiac arrhythmia. *Nature* **376**:683-685.

Benson DW, Silberbach GM, Kavanaugh-McHugh A, Cottrill C, Zhang Y, Riggs S, Smalls O, Johnson MC, Watson MS, Seidman JG, Seidman CE, Plowden J, Kugler JD. (1999) Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest*. **104**:1567-73.

Bers DM (2000) Calcium fluxes involved in control of cardiac myocyte contraction. *Circ Res*. **87**:275-281.

Berthier C, Blaineau S. (1997) Supramolecular organization of the subsarcolemmal cytoskeleton of adult skeletal muscle fibers. A review. *Biol Cell.* **89**:413-34.

Bezzina C, Veldkamp MW, van Den Berg MP, Postma AV, Rook MB, Viersma JW, van Langen IM, Tan-Sindhunata G, Bink-Boelkens MT, van Der Hout AH, Mannens MM, Wilde AA. (1999) A single Na(+) channel mutation causing both long-QT and Brugada syndromes. *Circ Res.* **85**:1206-1213.

Bhavsar PK, Brand NJ, Yacoub MH, Barton PJR (1996) Isolation and characterisation of human cardiac troponin I gene (*TNNI3*). *Genomics* **35**:11-23.

Bione S, D'Adamo P, Maestrinr E, Gedoen AK, Bolhuis PA, Toniolo D. (1996) A novel X-linked gene, G45, is responsible for Barth syndrome. *Nat. Genet*. **12**:385-389.

Blair E, Redwood C, Ashrafian H, Oliveira M, Broxholme J, Kerr B, Salmon A, Ostman-Smith I, Watkins H (2001) Mutations in the gamma (2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet* **10**:1215-1220.

Bleyl SB, Mumford BR, Thompson V, Carey JC, Pysher TJ, Chin TK, Ward K (1997) Neonatal lethal noncompaction of the left ventricular myocardium is allelic with Barth syndrome. *Am J Hum Genet* **61**:868-872.

Bockholt SM, Otey CA, Glenney JR Jr, Burridge K. (1992) Localization of a 215-kDa tyrosine-phosphorylated protein that cross-reacts with tensin antibodies. *Exp Cell Res*. **203**:39-46.

Bonne G, Carrier L, Richard P, Hsinque B, Schwartz K (1998) Familial hypertrophic cardiomyopathy from mutations to functional defects. *Circ Res*. **83**:580-593.

Bordo D. (1993) ENVIRON: a software package to compare protein three-dimensional structures with homologous sequences using local structural motifs. *Comput Appl Biosci*. **9**:639-645.

Brachfeld N, Gorlin R. (1959) Subaortic stenosis: a revised concept of the disease. *Medicine (Baltimore)***. 38**:415-33.

Brand T, Schneider MD. (1995) The TGF beta superfamily in myocardium: ligands, receptors, transduction, and function. *J Mol Cell Cardiol*. **27**:5-18.

Braunwald E. (2002) Hypertrophic cardiomyopathy-the benefits of a multidisciplinary approach. *N Engl J Med*. **347**:1306-1307.

Brenner SE, Chothia C, Hubbard TJ. (1998) Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships. *Proc Natl Acad Sci USA*. **95**:6073-6078.

Brent LB, Aburano A, Fisher DL, Moran TJ, Myers JD, Taylor WJ. (1960) Familial muscular subaortic stenosis: an unrecognized form of "idiopathic heart diseases," with clinical and autopsy observations. *Circulation*. **21**:167-80.

Bretscher A, Weber K. (1980) Fimbrin, a new microfilament-associated protein present in microvilli and other cell surface structures. *J Cell Biol*. **86**:335-40.

Brink AJ and Torrington M (1977) Progressive familial heart block-two type. *S Afr. Med. J* **52**:53-59.

Brink PA, Ferreira A, Moolman JC, Weymar H, Van der Merwe P-L, Corfield VA (1995) Gene for progressive familial heart block I maps to chromosome 19q13.3. *Circulation* **91:**1633-1640.

Brindle PK, Montminy MR. (1992) The CREB family of transcription activators. *Curr Opin Genet Dev*. **2**:199-204.

Brugada P, Brugada J. (1992) Right bundle branch block, persistent ST segment elevation and sudden death: A distinct clinical and electrocardiographic syndrome. *J. Am. Coll. Cardiol.* **20**:1391–1396.

Buratti R, Prestipino G, Menegazzi P, Treves S, Zorzato F. (1995) Calcium dependent activation of skeletal muscle Ca2+ release channel (ryanodine receptor) by calmodulin. *Biochem Biophys Res Commun*. **213**:1082-1090.

Carlson M. (1999) Glucose repression in yeast. *Curr Opin Microbiol*. **2**:202-207.

Carrier L, Hengstenberg C, Beckmann JS, Guicheney P, Dufour C, Bercovici J, Dausse E, Berebbi-Bertrand I, Wisnewsky C, Pulvenis D, *et al*. (1993) Mapping of a novel gene for familial hypertrophic cardiomyopathy to chromosome 11. *Nat Genet*. **4**:311-313.

Carrier L, Bonne G, Bahrend E, Yu B, Richard P, Neil F, Hainque B, Gautel M, Cruaud C, Gary F, Labeit S, Bouhour JB, Dbourg O, Desnos M, Hagege AA, Trent RJ, Komajda M, Schwartz K (1997) Organisation and sequence of human cardiac myosin binding protein C gene (*MYBPC3)* and identification of mutations predicted to produce truncated proteins in familial hypertrophic cardiomyopathy. *Circ Res*. **80**:427-434.

Cazorla O, Freiburg A, Helmes M, Centner T, McNabb M, Trombitas K, Labeit S, Granzier H. (2000) Differential expression of cardiac titin isoforms and modulation of cellular stiffness. *Circ Res* **86**:59-67.

Charron P, Dubourg O, Desnos M, Bennaceur M, Carrier L, Camproux AC, Isnard R, Hagege A, Langlard JM, Bonne G, Richard P, Hainque B, Bouhour JB, Schwartz K, Komajda M. (1998) Clinical features and prognostic implications of familial hypertrophic cardiomyopathy related to the cardiac myosin-binding protein C gene. *Circulation*. **97**:2230-2236.

Chen BPC, Wolfgang CD, Hai T. (1996) Analysis of ATF3, a transcription factor induced by physiological stresses and modulated by gadd153/Chop10. *Mol Cell Biol* **16**:1157–1168.

Chen CY, Schwartz RJ. (1996) Recruitment of the tinman homolog Nkx2.5 by serum response factor activates cardiac alpha-actin gene transcription. *Mol Cell Biol* **16**:6372-6383.

Chen Q, Kirsch GE, Zhang D, Brugada R, Brugada J, Brugada P, Potenza D, Moya A, Borggrefe M, Breithardt G, Ortiz-Lopez R, Wang Z, Antzelevitch C, O'Brien RE, Schulze-Bahr E, Keating MT, Towbin JA, Wang Q. (1998) Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* **392**:293-296.

Chu H-M, Tan Y, Kobierski LA, Balsam LB, and Comb MJ. (1994) Interactions among LRF-1, JunB, c-Jun, and c-Fos define a regulatory program in the G1 phase of liver regeneration. *Mol. Endocrinol.* **8**:59-68.

Collins FS, Morgan M, Patrinos A. (2003) The Human Genome Project: lessons from large-scale biology. *Science*. **300**:286-90.

Crabtree GR. (1999) Generic signals and specific outcomes: signaling through Ca^{2+}, calcineurin, and NF-AT. *Cell*. **96**:611-614.

Crabtree GR. (2001) Calcium, calcineurin, and the control of transcription. *J Biol Chem.* **276**:2313-2316.

Craig SW, Pardo JV. (1983) Gamma actin, spectrin, and intermediate filament proteins colocalize with vinculin at costameres, myofibril-to-sarcolemma attachment sites. *Cell Motil.* **3**:449-462.

D'Adamo P, Fassone L, Gedeon A, Janssen EA, Bione S, Bolhuis PA, Wilson M, Haan E, Orstavik KH, Patton MA, Green AJ, Zammarchi E, Donati MA, Toniolo D (1997) The X-linked gene *G4.5* is responsible for different infantile dilated cardiomyopathies. *Am. J. Hum. Genet*. **61**:868-872.

Dalla L, Hoffmann E, Floroff M, Jackowski G. (1989) Isolation and nucleotide sequence of the cDNA encoding human ventricular myosin light chain 2. *Nucleic Acids Res*. **17**:2360.

Dausse E, Komajda M, Fetler L, Dubourg O, Dufour C, Carrier L, Wisnewsky C, Bercovici J, Hengstenberg C, al-Mahdawi S, et al. (1993) Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and identification of a hot spot for mutations in the beta-myosin heavy chain gene. *J Clin Invest*. **92**:2807-2813.

Davies, M. J., Anderson, R. H., and Becker, A. E. (1983).The conduction system of the heart, London, Butterworths.

Deinum J, van Gool JM, Kofflard MJ, ten Cate FJ, Danser AH. (2001) Angiotensin II type 2 receptors and cardiac hypertrophy in women with hypertrophic cardiomyopathy. *Hypertension.* **38**:1278-1281.

de Lange WJ. (2004) An investigation of myosin binding protein C mutations in South Africa and a search for ligands binding to myosin binding protein C. Ph.D thesis, University of Stellenbosch.

De Meeus A, Stephan E, Debrus S, Jean MK, Weissenbach J, Demaille J, Bouvagnet P (1995) An isolated cardiac conduction disease maps to chromosome 19q. *Circ. Res*. **77:**735-740.

Depré C, Rider MH, Hue L (1998) Mechanism of control of heart glycolysis. *Eur J Biochem* **258**:277-290.

Deschenes I, Baroudi G, Berthet M, Barde I, Chalvidan T, Denjoy I, Guicheney P, Chahine M. (2000) Electrophysiological characterization of SCN5A mutations causing long QT (E1784K) and Brugada (R1512W and R1432G) syndromes. *Cardiovasc Res.* **46**:55-65.

Devereux RB, Alonso DR, Lutas EM, Gottlieb GJ, Campo E, Sachs I, Reichek N. (1986) Echocardiographic assessment of left ventricular hypertrophy: comparison to necropsy findings. *Am J Cardiol*.**57**:450-458.

DeWan AT, Parrado AR, Matise TC, Leal SM. (2002) The map problem: a comparison of genetic and sequence-based physical maps. *Am J Hum Genet*. **70**:101-107.

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380**:152-154.

Drysdale BE, Howard DL, Johnson RJ. (1996) Identification of a lipopolysaccharide inducible transcription factor in murine macrophages. *Mol Immunol*. **33**:989-998.

Dumaine R, *et al*. (1999) Ionic mechanisms responsible for the electrocardiographic phenotype of the Brugada syndrome are temperature dependent. *Circ. Res*. **85**:803-809.

Durocher D, Charron F, Warren R, Schwartz R, Nemer M. (1997) The cardiac transcription factors Nkx2.5 and GATA-4 are mutual cofactors. *EMBO J* **16**:5687-5696.

Elliott P (2000) Cardiomyopathy. Diagnosis and management of dilated cardiomyopathy. *Heart* **84**:106–112.

Erdmann J, Railble J, Maki-Abadi J, Hummel M, Hammann J, Wollnik B, Frantz E, Fleck E, Hetzer R, Regitz-Zagrosek V. (2001) Spectrum of clinical phenotypes and gene variants in cardiac myosin-binding protein C mutation carriers with hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.* **38**:322-330.

Fananapazir L, Epstein ND. (1994) Genotype-phenotype correlations in hypertrophic cardiomyopathy. Insights provided by comparisons of kindreds with distinct and identical beta-myosin heavy chain gene mutations. *Circulation.* **89**:22-32.

Farber JM. (1992) A collection of mRNA species that are inducible in the RAW 264.7 mouse macrophage cell line by gamma interferon and other agents *Mol Cell Biol.* **12**:1535-1545.

Fatkin D, Macrae C, Saski T, Wolff MR, Porcu M, Frenneaux M, Atherton J, Vidaillet HJ, Spudich S, De Girolami U, Seidman JG, Seidman CE (1999) Missense mutations in the rod

domain of the lamin A/C gene as cause of dilated cardiomyopathy and conduction system disease. *N. Engl. J.Med*. **341**:1715-1724.

Fatkin D, McConnell BK, Mudd JO, Semsarian C, Moskowitz IG, Schoen FJ, Giewat M, Seidman CE, Seidman JG. (2000) An abnormal Ca(2+) response in mutant sarcomere protein-mediated familial hypertrophic cardiomyopathy. *J Clin Invest*. **106**:1351-1359.

Fatkin D, Graham MR. (2002) Molecular mechanisms of inherited cardiomyopathies. *Physiol. Rev*. **82**: 945-980.

Fentzke RC, Korcarz CE, Lang RM, Lin H, Leiden JM. (1998) Dilated cardiomyopathy in transgenic mice expressing a dominant-negative CREB transcription factor in the heart. *J Clin Invest.* **101**:2415-2426.

Fernandez P, Corfield VA, Brink PA. (2004). Progressive familial heart block type II (PFHBII): a clinical profile from 1977 to 2003. *Cardiovasc J South Afr.* **15**:129-32.

Fernandez P. (2004) A candidate and novel gene search to identify the PFHBII-causative gene. Ph.D thesis, University of Stellenbosch.

Filatov VL, Katrukha AG, Bulargina TV, Gusev NB. (1999) Troponin: structure, properties, and mechanism of functioning. *Biochemistry (Mosc)***. 64**:969-85.

Fodor WL, Darras B, Seharaseyon J, Falkenthal S, Francke U, Vanin EF (1989) Human ventricular/slow twitch myosin alkali light chain gene: characterisation, sequence, and chromosomal location. *J Biol Chem*. **271**:27039-27043.

Foulkes NS, Borrelli E, Sassone-Corsi P. (1991) CREM gene: use of alternative DNA-binding domains generates multiple antagonists of cAMP-induced transcription. *Cell*. **64**:739-749.

Franz WM, Muller M, Muller OJ, Herrmann R, Rothmann T, Cremer M, Cohn Rd, Voit T, Katus HA (2000) Association of nonsense mutation of dystrophin gene with disruption of sarcoglycan complex in X-linked dilated cardiomyopathy. *Lancet* **355**:1781-1785.

Freiburg A, Trombitas K, Hell W, Cazorla, O, Fougerousse F, Centner T, Kolmerer B, Witt C, Beckman JS, Gregorio CC, Granzier H, Labeit S (2000) Series of exon-skipping events in the elastic spring region of titin as the structural basis for myofibrillar elastic diversity. *Circ Res* **36**:1114-1121.

Frey, N., Olson, E. N. (2003) Cardiac hypertrophy: the good, the bad, and the ugly. *Annu Rev Physiol*. **65**:45-79.

Fürst DP, Osborn M, Nave R, Weber K. (1988) The organisation of titin filaments in the half sarcomere revealed by monoclonal antibodies in immunoelectron microscopy: a map of ten nonrepititive epitopes starting at the Z-line extends to the M line. *J Cell Biol* **106**:1563-1572.

Fujita H, Sugiura S, Momomura S, Omata M, Sugi H, Sutoh K. (1997) Characterization of mutant myosins of Dictyostelium discoideum equivalent to human familial hypertrophic cardiomyopathy mutants. Molecular force level of mutant myosins may have a prognostic implication. *J Clin Invest*. **99**:1010-1015.

Fung DCY, Yu B, Littlejohn T, Trent RJ. (1999) An online locus-specific database for familial hypertrophic cardiomyopathy. *Hum Mutat* **14:**326-332.

Galperin MY. (2004) The Molecular Biology Database Collection: 2004 update. *Nucleic Acids Res*. **32** Database issue: D3-22**.**

Ganong, W. F. (1991). Physiology of nerve and muscle cells. Excitable tissue: Muscle.(16th ed.), Review of Medical Physiology (56-73). San Francisco: Prentice-Hall International Inc.

Gasser RB. (1997) Mutation scanning methods for the analysis of parasite genes. *Int J Parasitol*. **27:**1449-1463.

Gautel M, Zuffardi O, Freiburg A, Labeit S. (1995) Phosphorylation switches specific for cardiac isoform of myosin binding protein C: a modulator of cardiac contraction. *EMBO J* **14**: 1952-1960.

Geier C, Oezcelik C, perrot A, Bit-Avragm N Scheffold T, Osterziel KJ. (2001) Muscle LIM protein: a novel disease gene for hypertrophic cardiomyopathy? *Circulation* ***104 suppl*** **II**:II-521**.**

Gellen ME, George AL, Jr., Chen LQ, Chahine M, Horn R, Barchi RL, Kallen RG. (1992) Primay structure and functional expression of the human cardiac tetrodotoxin-insensitive voltage-dependent sodium channel. *Pro. Natl. Acad. Sci. USA*. **89**:554-558.

Genschel J, Schmidt HH. (2000) Mutation in the *LMNA* gene encoding lamin A/C. *Hum Mutat* **16**:451-459.

Gerace L. (1986) Nuclear lamina and organization of nuclear architecture. *Trends Biochem Sci* **11:**443-446**.**

Gerull B, Gramlich M, Atherton J, McNabb M, Trombitas K, Sasse-Klaassen S, Seidman CE, Granzier H, Labeit S, Frenneaux M, Thierfelder L. (2002) Mutations of *TTN*, encoding the giant muscle filament titin, cause familial dilated cardiomyopathy. *Nat genet* **30**:201-204.

Gillum RF. (1986) Idiopathc cardiomyopathy in the United States. *Am. Heart. J*. **111**:752-755.

Goldmuntz E, Geiger E, Benson DW. (2001) *NKX2.5* mutations in patients with tetralogy of fallot. *Circulation.* **104**:2565-2568.

Gollob MH, Green MS, Tang AS, *et al*. (2001) Identification of a gene responsible for familial Wolff-Parkinson-White syndrome. *N Engl J Med* **344:**1823-1831.

Gonzalez GA, Yamamoto KK, Fischer WH, Karr D, Menzel P, Biggs W 3rd, Vale WW, Montminy MR. (1989) A cluster of phosphorylation sites on the cyclic AMP-regulated nuclear factor CREB predicted by its sequence. *Nature*. **337**:749-752.

Granger BL, Lazarides E. (1978) The existence of an insoluble Z disc scaffold in chicken skeletal muscle. *Cell* **15**: 1253-68.

Grant AO, Carboni MP, Neplioueva V, Starmer CF, Memmi M, Napolitano C, Priori S. (2002) Long QT syndrome, Brugada syndrome, and conduction system disease are linked to a single sodium channel mutation. *J Clin Invest.* **110**: 1201-9.

Grünig E, Tasman JA, Kucherer H, Franz W, Kubler W, Katus HA. (1998) Frequency and phenotypes of familial dilated cardiomyopathy. *J. Am. Coll. Cardiol*. **31**: 186-194.

Grzeskowiak R, Witt H, Drungowski M, Thermann R, Hennig S, Perrot A, Osterziel KJ, Klingbiel D, Scheid S, Spang R, Lehrach H, Ruiz P. (2003) Expression profiling of human idiopathic dilated cardiomyopathy. *Cardiovasc Res.* **59**:400-411.

Hai TW, Liu F, Coukos WJ, Green MR. (1989) Transcription factor ATF cDNA clones: an extensive family of leucine zipper proteins able to selectively form DNA-binding heterodimers. *Genes Dev*. **3**:2083-2090.

Halford NG, Hardie DG. (1998) SNF1-related protein kinases: global regulators of carbon metabolism in plants? *Plant Mol Biol*. **37**:735-48.

Hallopeau L. (1869) Retrecissement ventriculo-aotique. *Gaz Med Pars* **24**: 683.

Harshfield GA, Grim CE, Hwang C, Savage DD, Anderson SJ. (1995) Genetic and environmental influences on echocardiographically determined left ventricular mass in black twins. *Am J Hypertens*. **3**:538-43.

Hasemann CA, Istuan ES, Uyeda K, Deisenhofer J. (1996) The crystal structure of the bifunctional enzyme 6-phosphfructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. *Structure* **4**:1017-1029.

Heine-Suner D, Diaz-Guillen MA, de Villena FP, Robledo M, Benifer J, de Cordobas R. (1997) A high resolution map of the regulator of the complement activation gene cluster on 1q32 that integrates new genes and markers. *Immunogenetics* **45**:422-427.

Heine-Suner D, Diaz-Guillen MA, Lange AJ, Rodriguez de Cardoba S. (1998) Sequence and structure of the human 6-phosphfructo-2-kinase/fructose-2,6-bisphosphatase heart isoform gene (*PFKFB2*). *Eur J Biochem* **254**:103-110.

Hirota H, Yoshida K, Kishimoto T, Taga T. (1995) Continuous activation of gp130, a signal-transducing receptor component for interleukin 6-related cytokines, causes myocardial hypertrophy in mice. *Proc Natl Acad Sci USA*. **92**:4862-4866.

Hodgkin AL, Huxley AF. (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**: 500–544.

Hoeffler JP, Meyer TE, Yun Y, Jameson JL, Habener JF. (1988) Cyclic AMP-responsive DNA-binding protein: structure based on a cloned placental cDNA. *Science.* **1242**:1430-1433.

Hoffman EP, Brown RH, Kunkel LM. (1987) Dystrophin: the protein product of the Duchenne muscular dystrophy locus. *Cell* **51**:919-928.

Homsher E, Lee DM, Morris C, Pavlov D, Tobacman LS. (2000) Regulation of force and unloaded sliding speed in single thin filaments: effects of regulatory proteins and calcium. *J Physiol*. **524**:233-243.

Hsu J-C, Bravo R, and Taub R. (1992) Interactions among LRF-1, JunB, c-Jun, and c-Fos define a regulatory program in the G1 phase of liver regeneration. *Mol. Cell. Biol.* **12**:4654-4665.

Huxley AF. (1957) Muscle structure and theories of contraction. *Prog Biophys Biophys Chem* **7**: 255-318.

Huxley HE, Brown W. (1967) The low-angle x-ray diagram of vertebrate striated muscle and its behaviour during contraction and rigor. *J Mol Biol.* **30**: 383-434.

Huxley HE. (1969) The mechanism of muscular contraction. *Science* **164**:1356-1366.

Huxley HE. (1974) Structural aspects of energy conversion in muscle. *Ann N Y Acad Sci*. **227**: 500-503.

Huynen M, Doerks T, Eisenhaber F, Orengo C, Sunyaev S, Yuan Y, Bork P. (1998) Homology-based fold predictions for Mycoplasma genitalium proteins. *J Mol Biol*. **280**:323-326.

Hwang JJ, Allen PD, Tseng GC, Lam CW, Fananapazir L, Dzau VJ, Liew CC. (2002) Microarray gene expression profiles in dilated and hypertrophic cardiomyopathic end-stage heart failure. *Physiol Genomics*. **10**:31-44.

Ikeda Y, Hiroi Y, Hosoda T, Utsunomiya T, Matsuo S, Ito T, Inoue J, Sumiyoshi T, Takano H, Nagai R, Komuro I. (2002) Novel point mutation in the cardiac transcription factor *CSX/NKX2.5* associated with congenital heart disease. *Circ J.* **66**:561-563.

International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**:860-921.

Ito N, Kagaya Y, Weinberg EO, Barry WH, Lorell BH. (1997) Endothelin and angiotensin II stimulation of Na+-H+ exchange is impaired in cardiac hypertrophy. *J Clin Invest*. **99**:125-135.

Itoh-Satoh M, Hayashi T, Nishi H, *et al*. (2002) Titin mutations as the molecular basis for dilated cardiomyopathy. *Biochem Biophys Res Commun.* **291**:385–393.

Jaenicke T, Diederich KW, Haas W, Schleich J, Lichter P, Pfordt M, Bach A, Vosberg HP. (1990) The complete sequence of the human β-myosin heavy chain gene and analysis of its product. *Genomics* **8**:194-206.

Jarcho JA, McKenna W, Pare JAP, Solomon SD, Holcombe RF, Dickie S, Levi T, Donis-Keller H, Seidman JG, Seidman CE. (1989) Mapping a gene for familial hypertrophic cardiomyopathy to chromosome 14q1. *N. Engl.J. Med* **321**:1372-1378.

Jeon JT, Carlborg O, Tornsten A, Giuffra E, Amarger V, Chardon P, Andersson-Eklund L, Andersson K, Hansson I, Lundstrom K, Andersson L. (1999) A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. *Nat Genet*. **21**:157-158.

Jing JP, Huang QQ, Yeh HI, Lin JJC. (1992) Complete nucleotide sequence and structural organisation of rat cardiac troponin T gene: a single gene generates embryonic and adult isoforms via developmentally regulated alternative splicing. *J Mol Biol*. **227**:1269-1276.

Jung D, Duclos F, Apostol B, Straub V, Lee CJ, Allamand V, Venzke PD, Sunada Y, Moomaw RC, Leveille JC, Slaughter AC, Crawford OT, Mcpherson DJ, Campell PK. (1996) Characterisation of δ-sarcoglycan, a novel component of the oligomeric sarcoglycan complex involved in Limb-Girdle Muscular dystrophy. *JBC* **271**:32321-32329.

Jung WI, Sieverding L, Breuer J, Hoess T, Widmaier S, Schmidt O, Bunse M, van Erckelens F, Apitz J, Lutz O, Dietze GJ. (1998) 31P NMR spectroscopy detects metabolic abnormalities in asymptomatic patients with hypertrophic cardiomyopathy. *Circulation*. **97**:2536-2542.

Kamisago M, Sharma S, DePalma SR, Solomon S, Sharma P, McDonough B, Smoot L, Mullen MP, Woolf PK, Wigle D, Seidman JG, Seidman CE. (2000) Mutations in sarcomere protein genes as a cause of dilated cardiomyopathy. *N Engl. J. Med*. **343:**1688-1696.

Karplus K, Barrett C, Hughey R. (1998) Hidden Markov models for detecting remote protein homologies. *Bioinformatics*. **14**:846-856.

Kasahara H, Lee B, Schott JJ, Benson DW, Seidman JG, Seidman CE, et al. (2000) Loss of function and inhibitory effects of human *CSX/NKX2.5* homoeprotein mutations associated with congenital heart disease. *J Clin Inves* **106**:299-308.

Kass S, MacRae C, Graber HL, Sparks E, McNamara D, Boudoulas H, Basson CT, Baker PB, Fishman MC, Cox N, Kong A, Wooley CF, Seidman JG, Seidman CE. (1994) A gene defect that causes conduction system disease and dilated cardiomyopathy maps to chromosome 1p1-1q1. *Nat. Genet*. **7**:546-551.

Kaufman SJ, Bielser D, Foster RF. (1990) Localization of anti-clathrin antibody in the sarcomere and sensitivity of myofibril structure to chloroquine suggest a role for clathrin in myofibril assembly. *Exp Cell Res.* **191**:227-38.

Kimura A, Harada H, Park JE, Nishi H, Satoh M, Takahashi M, Hiroi S, Sasaoka T, Ohbuchi N, Nakamura T, Koyanagi T, Hwang TH, Choo JA, Chung KS, Hasegawa A, Nagai R, Okazaki O, Nakamura H, Matsuzaki M, Sakamoto T, Toshima H, Koga Y, Imaizumi T, Sasazuki T. (1997) Mutation in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. *Nat.Genet*. **16**:379-382.

Kimura Y, Kurzydlowski K, Tada M, MacLennan DH. (1997) Phospholamban inhibitory function is activated by depolymerization. *J Biol Chem.* **272**:15061-15064.

Klues HG, Schiffers A, Maron BJ. (1995) Phenotypic spectrum and patterns of left ventricular hypertrophy in hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.* **26**:1699-1708.

Koenig M, Hoffman EP, Bertelson CJ, Monaco AP, Feener C, Kunkel LM. (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organisation of the DMD gene in normal and affected individuals. *Cell* **50**:509-517.

Kokado H, Shimizu M, Yoshio H, *et al*. (2000) Clinical features of hypertrophic cardiomyopathy caused by a Lys183 deletion mutation in cardiac troponin I gene. *Circulation* **102**:663-669.

Koppensteiner WA, Lackner P, Wiederstein M, Sippl MJ. (2000) Characterization of novel proteins based on known protein structures. *J Mol Biol.* **296**:1139-1152.

Koteliansky VE, Gneushev GN. (1983) Vinculin localization in cardiac muscle. *FEBS Lett.* **159**: 158-60.

Koteliansky V, Orgryzko E, Zhidkova N *et al*. (1992) An additional exon in the human vinculin gene specifically encodes meta-vinculin-specific different peptide: cross-species comparison reveals variable and conserved motifs in the meta-vinculin insert. *Eur J Biochem*. **204**:767-772.

Kunst G, Kress KR, Gruen M, Uttenweiler D, Gautel M, Fink RHA. (2000) Myosin binding protein C, a phosphorylation-dependent force regulator in muscle that controls the attachment of myosin heads by its interaction with myosin S2. *Circ Res* **86**:51-58.

Kurabayashi M, Komuro I, Tsuchimochi H, Takaku F, Yazaki Y. (1988) Molecular cloning and characterization of human atrial and ventricular myosin alkali light chain cDNA clones. *J Biol Chem*. **263**:13930-13936.

Lander ES, Schork NJ. (1994) Genetic dissection of complex traits. *Science* **265**:2037–2048.

Lang T, Yu L, Tu Q, Jiang J, Chen Z, Xin Y, Liu G, Zhao S. (2000) Molecular cloning, genomic organization and mapping of *PRKAG2*, a heart abundant gamma2 subunit of 5'-AMP-activated protein kinase, to human chromosome 7q36. *Genomics* **70**:258-263.

Lankford EB, Epstein ND, Fananapazir L, Sweeney HL (1995) Abnormal contractile properties of muscle fibers expressing b-myosin heavy chain gene mutations in patients with hypertrophic cardiomyopathy. *J Clin Invest* **95**: 1409-1414.

Laufs U, Kilter H, Konkol C, Wassmann S, Bohm M, Nickenig G. (2002) Impact of HMG CoA reductase inhibition on small GTPases in the heart. *Cardiovasc Res*. **53**:911-920.

Lazarides E. (1978) The distribution of desmin (100 A) filaments in primary cultures of embryonic chick cardiac cells. *Exp Cell Res*. **112**:265-73.

Lechin M, Quinones MA, Omram A, Hill R, Yu Q-T, Rakowski W, Liew CC, Sole M, Roberts R, Marian AJ (1995) Angiotensin-I converting enzyme genotypes and left ventricular hypertrophic cardiomyopathy. *Circulation* **92**:1808-1812.

Leiden JM. (1997) The genetics of dilated cardiomyopathy--emerging clues to the puzzle. *N Engl J Med.* **337**:1080-1081.

Lenègre J, Moreau PH. (1963) Le bloc auriculo-ventriculaire chronique, Étude anatomique, clinique et histologique. *Arch Mal Coeur* **56**:867–888.

Lev M, Kinare SG, Pick A. (1970) The pathogenesis of atrioventricular block in coronary disease. *Circulation* **42**:409–425.

Li ZL, Lilienbaum A, Butler-Browne G, Paulin D (1989) Human desmin-coding gene: nucleotide sequence, characterisation, and regulation of expression during myogenesis and development. *Gene* **78**:243-254.

Li D, Tapscoft T, Gonzalez O, Burch PE, Quinones MA, Zoghbi WA, Hill R, Bachinski LL, Mann DL, Roberts R (1999) Desmin mutation responsible for idiopathic dilated cardiomyopathy. *Circulation* **100**:461-464.

Liang G, Wolfgang CD, Chen BP, Chen TH, Hai T. (1996) ATF3 gene. Genomic organization, promoter, and regulation. *J Biol Chem.* **271**:1695-1701.

Lim DS, Roberts R, Marian AJ. (2001) Expression profiling of cardiac genes in human hypertrophic cardiomyopathy: insight into the pathogenesis of phenotypes. *J Am Coll Cardiol.* **38**:1175-1180.

Lin D, Bobkova A, Homsher E, and Tobacman LS. (1996) Altered cardiac troponin T in vitro function in the presence of a mutation implicated in familial hypertrophic cardiomyopathy. *J Clin Invest* **97**:2842-2848.

Lin F, Worman HJ (1993) Structural organisation of the human gene encoding nuclear lamin A and nuclear lamin C. *J. Biol. Chem*. **268**:16321-16326.

Lowey S. (2002) Functional consequences of mutations in the myosin heavy chain at sites implicated in familial hypertrophic cardiomyopathy. *Trends Cardiovasc. Med.* **12**:348–354.

Lymn RW, Taylor EW (1971) Mechanism of adenosine triphosphate hydrolysis by actomyosin. *Biochemistry* **10**:4617-4624.

MacLeod AR, Gooding C. (1988) Human hTM alpha gene: expression in muscle and nonmuscle tissue. *Mol Cell Biol*. **8**:433-440.

Mamane Y, Heylbroeck C, Genin P, Algarte M, Servant MJ, LePage C, DeLuca C, Kwon H, Lin R, Hiscott J. (1999) Interferon regulatory factors: the next generation. *Gene*. **237**:1-14.

Marian AJ, Yu QT, Workman R, Greve G, Roberts R. (1993) Angiotensin-converting enzyme polyrmophism in hypertrophic cardiomyopathy and sudden cardiac death. *Lancet* **342**:1085-1086.

Marian AJ, Mares A, Kelly DP, Yu Q-T, Abchee AB, Hill R, Roberts R. (1995a) Sudden cardiac death in hypertrophic cardiomyopathy. Variability in phenotypic expression of β-myosin heavy chain mutations. *Eur Heart J* **16**:368-376.

Marian AJ, Roberts R. (1995) Recent advances in the molecular genetics of hypertrophic cardiomyopathy. *Circulation* **92:**1336-1347.

Maron BJ, Gottdiener JS, Roberts WC, et al. (1978) Left ventricular outflow tract obstruction due to systolic anterior motion of the anterior mitral leaflets in patients with concentric left ventricular hypertrophy. *Circualtion* **57**:527-533.

Maron BJ, Gardin JM, Gidding SS, Kurosaki TT, Bild DE. (1995) Prevalence of hypertrophic cardiomyopathy in a general population of young adults: echocardiographic analysis of 4111 subjects in the CARDIA Study-Coronary Artery Risk Development in (Young) Adults. *Circulation* **92**:785-789.

Maron BJ, Olivotto I, Spirito P, *et al*. (2000) Epidemiology of hypertrophic cardiomyopathy-related death: revisited in a large non-referral-based patient population. *Circulation* **102**:858-864.

McDonald KS, Moss RL. (2000) Strongly binding myosin crossbridges regulate loaded shortening and power output in cardiac myocytes. *Circ Res* **87**:768-773.

McGregor A, Blanchard AD, Rowe AJ, Critchley DR. (1994) Identification of the vinculin-binding site in the cytoskeletal protein alpha-actinin. *Biochem J.* **301**:225-33.

McTiernan CF, Frye SC, Lemster BH, Kinder EA, Ogletree-Hughes ML, Moravec CS, Feldman MA. (1999) The human phospholamban gene: structure and axpression. *J. Mol. Cell. Cardiol.* **31**:679-692.

Messina DA, Lemanski LF. (1989) Immunocytochemical studies of spectrin in hamster cardiac tissue. *Cell Motil Cytoskeleton*. **12**:139-49.

Meyer TE, Habener JF. (1993) Cyclic adenosine 3',5'-monophosphate response element binding protein (CREB) and related transcription-activating deoxyribonucleic acid-binding proteins. *Endocr Rev.* **14**:269-290.

Michele DE, Albayya FP, Metzger JM. (1999) Direct, convergent hypersensitivity of calcium-activated force generation produced by hypertrophic cardiomyopathy mutant alpha-tropomyosins in adult cardiac myocytes. *Nat Med.* **5**:1413-1417.

Michels VV, Moll PP, Miller FA, Tajik AJ, Chu JS, Driscoll DJ, Burnett JC, Rodeheffer RJ, Chesebro JH, Tazelaar HD. (1992) The frequency of familial dilated cardiomyopathy in a series of patients with idiopathic dilated cardiomyopathy. *N Engl J Med* **326**:77-82.

Mogensen J, Klausen IC, Pedersen AK, Egebkad G, Bross P, Kruse TA, Gregersen N, Hansen PS, Baandrup U, Borglum AD. (1999) α-Cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. *J. Clin. Invest*. **103**:R39-43.

Molkentin JD, Lu JR, Antos CL, Markham B, Richardson J, Robbins J, Grant SR, Olson EN. (1998) A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell*. **93**:215-28.

Moolman JC, Brink PA, Corfield VA. (1993) Identification of a new missense mutation at Arg403, a CpG mutation hotspot, in exon 13 of the beta-myosin heavy chain gene in hypertrophic cardiomyopathy. *Hum Mol Genet*. **2**:1731-1732.

Moolman JC, Brink PA, Corfield VA. (1995) Identification of a novel Ala797Thr mutation in exon 21 of the beta-myosin heavy chain gene in hypertrophic cardiomyopathy. *Hum Mutat*. **6**:197-198.

Moolman JC, Corfield VA, posen B, Ngumbela K, Seidman CE, Brink PA and Watkins H. (1997) Sudden death due to troponin T mutations. *J Am Coll Cardiol* **29**:549-555.

Moolman-Smook JC, Mayosi B, Brink P, Corfield VA. (1998) Identification of a new missense mutation in MyBP-C associated with hypertrophic cardiomyopathy. *J Med Genet*. **35:**253-254.

Moolman-Smook JC, De Lange WJ, Bruwer ECD, Brink PA, Corfield VA. (1999) The origins of hypertrophic cardiomyopathy-causing mutations in two South African sub-populations: a unique profile of both independent and founder events. *Am. J. Hum. Genet*. **65**:1308-1320.

Moolman-Smook JC, De Lange J, Brink A, Corfield A. (2000) Hypertrophic cardiomyopathy repealing tenets in South Africa. *Cardiovasc J S Afr*. **11:**202-209.

Moolman-Smook JC, Mayosi BM, Brink PA, Corfield VA. (2003) Molecular genetics of cardiomyopathy: changing times, shifting paradigms. *Cardiovasc J S Afr*. **14**:145-55.

Moorman AFM, de Jong F, Denyn MMFJ, Lamers WH. (1998) Development of the conduction system. *Circ Res*. **82**:629-644.

Morimoto S, Yanaga F, Minakami R, Ohtsuki I. (1998) Ca2+-sensitizing effects of the mutations at Ile-79 and Arg-92 of troponin T in hypertrophic cardiomyopathy. *Am J Physiol.* **275**:C200-207.

Muntoni F, Cau M, Congiu R, *et al*. (1993) Deletion of the dystrophin muscle-promoter region associated with X-linked dilated cardiomyopathy. *N. Engl. J. Med.* **329**:921-923.

Naccarelli VG, Antzelevitch C, Wolbrette DL, Luck JC. (2002) The Brugada syndrome. *Curr Opin Cardiol* **17**: 19–23.

Nakajima-Taniguchi C, Matsui H, Fujio Y, et al. (1997) Novel missense mutation in cardiac troponin T gene found in Japanese patient with hypertrophic cardiomyopathy. *J Mol Cell Cardiol.* **29**:839–843.

Nelson WJ, Lazarides E. (1983) Expression of the beta subunit of spectrin in nonerythroid cells. *Proc Natl Acad Sci U S A.* **80**:363-7.

Nezer C, Moreau L, Brouwers B, Coppieters W, Detilleux J, Hanset R, Karim L, Kvasz A, Leroy P, Georges M. (1999) An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nat Genet*. **21**:155-156.

Nigro V, de Sa Moreira e, Piluso G, Vainzof M, Belsito A, Plitano L, Puca AA, Passos-Bueno MR, Zatz M. (1996) Autosomal recessive limb-girdle muscular dystrophy, LGMD2F, is caused by a mutation in the delta-sarcoglycan gene. *Nat Genet* **14**:195-198.

Niimura H, Bachinski LL, Sangwatanaroj S, Watkins H, Chudley AE, McKenna W, Kristinsson A, Roberts R, Sole M, Maron BJ, Seidman JG, Seidman CE. (1998) Mutations in the gene for human cardiac myosin-binding protein C and late-onset familial hypertrophic cardiomyopathy. *N Engl J Med*. **338**:1248-1257.

Okamoto Y, Chaves A, Chen J, Kelley R, Jones K, Weed HG, Gardner KL, Gangi L, Yamaguchi M, Klomkleaw W, Nakayama T, Hamlin RL, Carnes C, Altschuld R, Bauer J, Hai T. (2001) Transgenic mice with cardiac-specific expression of activating transcription factor 3, a stress-inducible gene, have conduction abnormalities and contractile dysfunction. *Am J Pathol.* **159**:639-650.

Olson TM, Michels VV, Thibodeau SN, Tai Y-S, Keating MT. (1998) Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* **280:**750-752.

Olson TM, Doan TP, Kishimoto NY, Whitby FG, Ackerman MJ, Fananapazir L (2000) Inherited and de novo mutations in the cardiac actin gene cause hypertrophic cardiomyopathy. *J Mol Cell Cardiol* **32**:1687-1694.

Olson TM, Kishimoto NY, Whitby FG, Michels VV (2001) Mutations that alter the surface charge of alpha-tropomyosin are associated with dilated cardiomyopathy. *J Mol Cell Cardiol* **33**:723-732.

Olson TM, Illenberger S, Kishimoto YN, Huttelmaier S, Keating TM, Jockusch MB (2002) Metavinculin mutations alter actin interaction in dilated cardiomyopathy. *Circulation* **105**:431-437.

Opie, L. H., (1991). Heart cells and organelles. Swynghedauw B. and Taegtmeyer H. (2nd ed.), The heart: Physiology and metabolism (43-68). New York: Raven Press.

Orita M, Suzuki Y, Sekiya T, Hayashi K. (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics*. **5**:874-879.

Osoegawa K, Mammoser AG, Wu C, Frengen E, Zeng C, Catanese JJ, de Jong PJ. (2001) A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res*. **11**:483-496.

Osterop AP, Kofflard MJ, Sandkuijl LA, et al. (1998) AT1 receptor A/C1166 polymorphism contributes to cardiac hypertrophy in subjects with hypertrophic cardiomyopathy. *Hypertension* **32**:825-830.

Pardo JV, Siliciano JD, Craig SW. (1983) Vinculin is a component of an extensive network of myofibril-sarcolemma attachment regions in cardiac muscle fibers. *J Cell Biol*. **97**:1081-1088.

Paré JAP, Fraser RG, Pirozynski WJ, Shanks JA, Stubington D. (1961) Hereditary cardiovascular dysplasia: A form of familial cardiomyopathy. *Am J Med* **31**:37-62.

Park J, Karplus K, Barrett C, Hughey R, Haussler D, Hubbard T, Chothia C. (1998) Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. *J Mol Biol*. **284**:1201-1210.

Park KY, Dalakas MC, Goebel HH, Ferrans VJ, Semino-Mora C, Litvak S, Takeda K, Goldfarb LG. (2000) Desmin splice variants causing cardiac and skeletal myopathy. *J Med Genet*. **37**:851-857.

Papassotiropoulos A, Hock C, Nitsch RM. (2001) Genetics of interleukin 6: implications for Alzheimer's disease. *Neurobiol Aging* **22**:863-871.

Patel R, Lim DS, Reddy D, *et al* (2000) Variants of trophic factors and expression of cardiac hypertrophy in patients with hypertrophic cardiomyopathy. *J Mol Cell Cardiol* **32**:2369-2377.

Pavalko FM, Otey CA, Simon KO, Burridge K. (1991) Alpha-actinin: a direct link between actin and integrins. *Biochem Soc Trans***. 19**:1065-9.

Pearson WR. (1996) Effective protein sequence comparison. *Methods Enzymol*. **266**:227-258.

Petrich BG, Wang Y. (2004) Stress-activated MAP kinases in cardiac remodeling and heart failure; new insights from transgenic studies. *Trends Cardiovasc Med*. **14**:50-55.

Pfeufer A, Osterziel KJ, Urata H, *et al.* (1996) Angiotensin-converting enzyme and heart chymase gene polymorphisms in hypertrophic cardiomyopathy. *J Am Coll Cardiol* **78**:362-364.

Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Seller JR, Fananapazir L, Epstein ND (1996) Mutation in either the essential or regulatory light chain of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat. Genet*. **13**:63-69.

Pola R, Flex A, Gaetani E, Dal Lago A, Gerardino L, Pola P, Bernabei R. (2002) The -174 G/C polymorphism of the interleukin-6 gene promoter is associated with Alzheimer's disease in an Italian population. *Neuroreport* **13**:1645-1647.

Posen BM, Moolman JC, Corfield VA, Brink PA. (1995) Clinical and prognostic evaluation of familial hypertrophic cardiomyopathy in two South African families with different cardiac beta myosin heavy chain gene mutations. Br *Heart J*. **74**:40-46.

Price M, Sanger JW. (1979) Intermediate filaments connect z-discs in adult chicken muscle. *J Exp Zool*. **208**:263-269.

Priori SG, Barhanin J, Hauer RN, Haverkamp W, Jongsma HJ, Kleber AG, McKenna WJ, Roden DM, Rudy Y, Schwartz K, et al. (1999a) Genetic and molecular basis of cardiac arrhythmias: Impact on clinical management (Part I and II). *Circulation* **99**:518-528.

Raharjo WH, Enarson P, Sullivan T, Stewart CL, Burke B. (2001) Nuclear envelope defects associated with LMNA mutations cause dilated cardiomyopathy and Emery-Dreifuss muscular dystrophy. *J Cell Sci*. **114**: 4447-57.

Randle PJ, Garland PB, Hales CN, Newsholme EA. (1963) The glucose fatty-acid cycle. Its role in insulin sensitivity and the metabolic disturbances of diabetes mellitus. *Lancet.* **1**:785-789.

Rayment I, Holden HM, Whittaker M, Yohn CB, Lorenz M, Holmes KC, Milligan RA (1993) Structure of the actin-myosin complex and its implications for muscle contraction. *Science* **261**:58-65.

Rayment I, Holden HM, Sellers JR, Fananapazir L, Epstein ND (1995) Structural interpretation of the mutations in the β-cardiac myosin that have been implicated in familial hypertrophic cardiomyopathy. *Proc Natl acad Sci USA* **92**: 3864-3868.

Redwood CS, Moolman-Smook JC, and Watkins H (1999) Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc.Res*. **44**:20-36.

Reedy MC. (2000) Visualizing myosin's power stroke in muscle contraction. *J Cell Sci*. **113**:3551-62.

Regitz-Zagrosek V, Erdmann J, Wellnhofer E, Raible J, Fleck E (2000) Novel mutation in the alpha-tropomyosin gene and transition from hypertrophic to hypocontractile dilated cardiomyopathy. *Circulation* **102**:E112-116.

Repasky EA, Granger BL, Lazarides E. (1982) Widespread occurrence of avian spectrin in nonerythroid cells. *Cell.* **29**:821-833.

Report of the World Health Organisation/International Society and Federation of Cardiology Task Force on the definition and classification of cardiomyopathies. (1980) *Circulation* **93**:841-842.

Report of the 1995 World Health Organisation/International Society and Federation of Cardiology Task Force. (1996) The definition and classification of cardiomyopathies. *Circulation* **93**:841-842.

Risch N, Merikangas K. (1996) The future of genetic studies of complex human diseases. *Science* **273**:1516**–**1517.

Roberts S L, Ervasti J M, Anderson R D, Ohlendieck K, Kahl S D, Zoloto D, Campbell K P. (1993) Disruption of the dystrophin-glycoprotein complex in the cardiomyopathic hamsters. *J. Biol. Chem.* **268**:11496-11499.

Roesler WJ, Vandenbark GR, Hanson RW. (1988) Cyclic AMP and the induction of eukaryotic gene transcription. *J Biol Chem*. **263**:9063-9066.

Ross OA, O'Neill C, Rea IM, Lynch T, Gosal D, Wallace A, Curran MD, Middleton D, Gibson JM. (2004) Functional promoter region polymorphism of the proinflammatory chemokine IL-8 gene associates with Parkinson's disease in the Irish. *Hum Immunol.* **65**:340-346.

Rust EM, Albayya RP and Metzger JM (1999) Identification of a contractile deficit in adult cardiac myocytes expressing hypertrophic cardiomyopathy-associated mutant troponin T proteins. *J. Clin. Invest*. **103**:1459-1467.

Sanbe A, Gulick J, Haye E, Warshaw D, Osinska H, Chan CB, Klevitsky R, Robbins J (2000) Myosin light chain replacement in the heart. *Am J Physiol Heart Circ Physiol* **279**:H1355-H1364.

Sata M, and Ikebe M. (1996) Functional analysis of the mutations in the human cardiac β-myosin that are responsible for familial hypertrophic cardiomyopathy. *J Clin Invest* **98**:2866-2873.

Satoh M, Takahashi M, Sakamoto T, Hiroe M, Marumo F, Kimura A (1999) Structural analysis of the titin gene in hypertrophic cardiomyopathy: Identification of a novel disease gene. *Biochem. Biophy. Res. Commun* **262**:411-217.

Schmitt JP, Kamisago M, Asahi M, Li HG, Ahmad F, Mende U, Kranias GE, MacLennan DH, Seidman GJ, Seidman EC (2003) Dilated cardiomyopathy and heart failure caused by a mutation in phospholamban. *Science* **299:**1410-1413.

Schott J-J, Alshinawi C, Kyndt F, Probst V, Hoorntje TM, Hulsbeek M, Wilde AA, Escande D, Mannens MM, Le Marec H (1999) Cardiac conduction defects associated with mutation in SCN5A. *Nat. Genet*. **23**: 20-21.

Schunkert H, Brockel U, Hengstenberg C, Luchner A, Muscholl MW, Kurzidim K, Kuch B, Doring A, Riegger GA, Hense HW. (1999) Familial predisposition of left ventricular hypertrophy. *J Am Coll Cardiol*. **33**:1685-1691.

Schwartz PJ, Priori SG, Locati E, Napolitano C, Cantum F, Towbin JA, Keating MT, Hammoude H, Brown AM, Chen LSK, Colatsky TJ. (1995) Long QT syndrome patients with mutations of the SCN5A and HERG genes have differential responses to Na^+ channel blockade and to increases in heart rate. *Circulation* **92**:3381–3386.

Schwartz PJ, Moss AJ, Priori SG, Wang Q, Lehmann MH, Timothy KW, Denjoy IF, Haverkamp W, Guicheney P, Paganini V, Scheinmann MM, Kames PS. (1997) Gene-specific influence on the triggers for cardiac arrest in the long QT syndrome. *Circulation* **96** (suppl I):I-212.

Sebkhi A, Zhao L, Lu L, Haley CS, Nunez DJ, Wilkins MR. (1999) Genetic determination of cardiac mass in normotensive rats: results from an F344xWKY cross. *Hypertension.* **33**:949-953.

Seidman JG, Seidman CE (2001) The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell* **104**:557-567.

Semple CA, Evans KL, Porteous DJ. (2001) Twin peaks: the draft human genome sequence. *Genome Biol*. **2**:COMMENT2003.

Semsarian C, Healey MJ, Fatkin D, Giewat M, Duffy C, Seidman CE, Seidman JG. (2001) A polymorphic modifier gene alters the hypertrophic response in a murine model of familial hypertrophic cardiomyopathy. *J Mol Cell Cardiol*. **33**:2055-2060.

Shear CR, Bloch RJ. (1985) Vinculin in subsarcolemmal densities in chicken skeletal muscle: localization and relationship to intracellular and extracellular structures. *J. Cell Biol.* **101**: 240-256.

Sheffield VC, Beck JS, Kwitek AE, Sandstrom DW, Stone EM. (1993) The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* **16:**325-332.

Siu BL, Niimura H, Osborne JA, Fatkin D, MacRae C, Solomon S, Benson W, Seidman JG, Seidman CE (1999) Familial dilated cardiomyopathy locus maps to chromosome 2q31. *Circulation* **99**: 1022-1026.

Solaro RJ, Van Eyk J (1996) Altered interaction among thin filament proteins modulate cardiac function. *J Mol Cell Cardiol*. **28**:217-230.

Song HY, Régnier CH, Kirschning CJ, Goeddel DV, Rothe M. (1997) Tumor necrosis factor (TNF)-mediated kinase cascades: bifurcation of nuclear factor-kappaB and c-jun N-terminal kinase (JNK/SAPK) pathways at TNF receptor-associated factor 2. *Proc. Natl. Acad. Sci. USA* **94**:9792–9796.

Spielman RS, McGinnis RE, Ewens WJ. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* **52**:506–513.

Spindler M, Saupe KW, Christe ME, Sweeney HL, Seidman CE, Seidman JG, Ingwall JS (1998) Diastolic dysfunction and altered energetics in the αMHC/+ mouse model of familial hypertrophic cardiomyopathy. *J. Clin. Invest.* **10**:1775-1783.

Stromer MH. (1995) Immunocytochemistry of the muscle cell cytoskeleton *Microsc. Res Techn* **31:** 95-105.

Stuurman N, Heins S, Aebi U. (1998) Nuclear lamins: their structure, assembly and interactions. *J. Struct. Biol*. **122**:42-66.

Sweeney HL, Feng HS, Yang Z, Watkins H (1998) Functional analysis of troponin T mutations that causes hypertrophic cardiomyopathy: insight into disease pathogenesis and troponin function. Proc. *Natl Acad. Sci. USA* **95**:14406-14410.

Tan HI, Bink-Boelkens MTE, Bezzina CR, Viswanathan PC, Beaufort-Krols GCM, Van Tintelen PJ, Van Den Berg, MP, Wilde AAM, Balser JR (2001) A sodium channel causes isolated cardiac conduction disorder. *Nature* **409**:1043-47.

Tanaka M, Kasahara H, Bartunkova S, Schinke M, Komuro I, Inagaki H, et al. (1998) Vertebrate homologs of tinman and bagpipe: Roles of the homeobox genes in cardiovascular development. *Dev Genet* **22**:239-249.

Tardiff JC, Hewett TE, Palmer BM, Olsson C, Factor SM, Moore RL, Robbins J, Leinwand LA. (1999) Cardiac troponin T mutations result in allele-specific phenotypes in a mouse model for hypertrophic cardiomyopathy. *J Clin Invest*. **104**:469-481.

Teare RD. (1958) Asymmetrical hypertrophy of the heart in young adults. *Br Heart J* **20**:1-8.

Tesson F, Dufour C, Moolman JC, Carrier L, al-Mahdawi S, Chojnowska L, Dubourg O, Soubrier E, Brink P, Komajda M, Guicheney P, Schwartz K, Feingold J. (1997) The influence of the angiotensin I converting enzyme genotype in familial hypertrophic cardiomyopathy varies with the disease gene mutation. *J Mol Cell Cardiol*. **29**:831-238.

Thierfelder L, MacRae C, Watkins H, Tomfohrde J, Williams M, McKenna W, Bohm K, Noeske G, Schlepper M, Bowcock A (1993) A familial hypertrophic cardiomyopathy locus maps to chromosome 15q2. *Proc. Natl. Acad.Sci USA* **90**:266-267.

Towbin JA, Hejtmancik JF, Brink P, Gelb B, Zhu XM, Chamberlain JS, McCarbe ERB, Swift M (1993) X-linked dilated cardiomyopathy. Molecular genetic evidence of linkage to

the Duchenne muscular dystrophy (dystrophin) gene at the Xq21 locus. *Circulation* **87**:1854-1865.

Towbin JA, Bowles KR, Bowles NE. (1999) Etiologies of cardiomyopathy and heart failure. *Nat Med*. **5**: 266-267.

Towbin JA. (2000) Molecular genetics of hypertrophic cardiomyopathy. *Curr Cardiol Rep*. **2**: 134-40.

Trinick J, Lowey S. (1977) M-protein from chicken pectoralis muscle: isolation and characterisation. *J Mol Biol.* **113**:343-368.

Tsubata S, Bowles KR, Vatta M, Zintz C, Titus J, Muhonen L, Bowles NE, Towbin JA (2000) Mutations in the human delta-sarcoglycan gene in familial and sporadic dilated cardiomyopathy. *J.Clin. Invest.* **106**:655-662.

Turner CE, Glenney JR Jr, Burridge K. (1990) Paxillin: a new vinculin-binding protein present in focal adhesions. *J Cell Biol*. **111**: 1059-68.

Turner NA, Xia F, Azhar G, Zhang X, Liu L, Wei JY. (1998) Oxidative stress induces DNA fragmentation and caspase activation via the c-Jun NH2-terminal kinase pathway in H9c2 cardiac muscle cells. *J Mol Cell Cardiol*. **30**:1789-1801.

Tyska MJ, Hayes E, Giewat M, Seidman CE, Seidman JG, and Warshaw DM. (2000) Single-molecule mechanics of R403Q cardiac myosin isolated from the mouse model of familial hypertrophic cardiomyopathy. *Circ Res* **86**:737-744.

Veldkamp MW, *et al.* (2000) Two distinct congenital errhythmias evoked by a multidysfunctional Na^+ channel. *Circ. Res*. **86**:e91-e97.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al. (2001) The sequence of the human genome. *Science* **291**: 1304-1351.

Wahl M. (1999) The heart is a muscle, too. *Quest* **6.**

Wallace DC. (2001) A mitochondrial paradigm for degenerative diseases and ageing. *Novartis Found Symp*. **235**:247-63.

Wang Q, Shen J, Splawski I, Atkinson D, Li Z, Robinson JL, Moss AJ, Towbin JA, Keating MT. (1995) SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell*. **80**: 805-11.

Wang Q, Li Z, Shen J, Keating MT. (1995) Genomic organization of the human *SCN5A* gene encoding the cardiac sodium channel. *Genomics* **34**:9-16.

Wannenburg T, Heijne GH, Geerdink JH, Van Den Dool HW, Janssen PML, De Tombe PP. (2000) Cross-bridge kinetics in rat myocardium: effect of sarcomere length and calcium activation. *Am J Physiol Heart Circ Physiol* **279**:H779-790.

Watkins H, Rosenzweig A, Hwang D-S, *et al*. (1992) Characteristics and prognostic implications of myosin missense mutations in familial hypertrophic cardiomyopathy. *N Engl J Med* **326**:1108-1114.

Watkins H, Thierfelder L, Anan R, Jarcho J, Matsumori A, McKenna W, Seidman JG, Seidman CE. (1993) Independent origin of identical beta cardiac myosin heavy-chain mutations in hypertrophic cardiomyopathy. *Am J Hum Genet*. **53**:1180-1185.

Watkins H, McKenna WJ, Thierfelder L, *et al.* (1995) Mutations in the genes for cardiac troponin T and alpha-tropomyosin in hypertrophic cardiomyopathy. *N Engl J Med* **332:**1058-1064.

Watkins H, Conner D, Thierfelder L, Jarcho JA, MacRae C, McKenna W, Maron BJ, Seidman JG, and Seidman CE (1995) Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. *Nat Genet* **11:**434-437.

Watkins H, Anan R, Coviello DA, Spirito P, Seidman JG, Seidman CE (1995) A de novo mutation in alpha-tropomyosin that cause hypertrophic cardiomyopathy. *Circulation* **91**:2302-2305.

Watkins H, Seidman CE, Seidman JG, Feng HS, Sweeney HL (1996) Expression and functional assessment of a truncated cardiac troponini T that cause hypertrophic cardiomyopathy: evidence for a dominant negative action. *J. Clin Invest*. **11**:2456-2461.

Watkins H. (2003) Genetic clues to disease pathways in hypertrophic and dilated cardiomyopathies. *Circulation*. **107**:1344-1346.

Weisberg A, Winegrad S. (1998) Relation between crossbridge structure and actomyosin ATPase activity in rat heart. *Circ Res* **83**:60-72.

Whiting A, Wardale J, Trinick J. (1989) Does titin regulate the length of muscle thick filaments? *J Mol Biol* **205**:163-169.

Wigle ED, Sasson Z, Henderson MA, Ruddy TD, Fulop J, Rakowski H, Williams WG. (1985) Hypertrophic cardiomyopathy. The importance of the site and the extent of hypertrophy. A review. *Prog Cardiovasc Dis*. **28**:1-83.

Yamada Y, Ichihara S, Fujimura T, Yokota M. (1997) Lack of association of polymorphisms of the angiotensin converting enzyme and angiotensinogen genes with nonfamilial hypertrophic or dilated cardiomyopathy. *Am J Hypertens*. **10**:921-928.

Yang J, Moravec CS, Sussman MA, DiPaola NR, Fu D, Hawthorn L, Mitchell CA, Young JB, Francis GS, McCarthy PM, Bond M. (2000) Decreased SLIM1 expression and increased gelsolin expression in failing human hearts measured by high-density oligonucleotide arrays. *Circulation* **102**:3046-3052.

Yin T, Sandhu G, Wplfgang CD, Burrier A, Webb RL, Rigel DF, Hai T, Whelan J. (1997) Tissue-specific pattern of stress kinase activation in ischemic/reperfused heart and kidney. *J. Biol. Chem.* **272**:19943–19950.

Yokoyama T, Nakano M, Bednarczyk JL, McIntyre BW, Entman M, Mann DL. (1997) Tumor necrosis factor-alpha provokes a hypertrophic growth response in adult cardiac myocytes. *Circulation*. **95**:1247-1252.

Yoneya K, Okamoto H, Machida M, Onozuka H, Noguchi M, Mikami T, Kawaguchi H, Murakami M, Uede T, Kitabatake A. (1995) Angiotensin-converting enzyme gene polymorphism in Japanese patients with hypertrophic cardiomyopathy. *Am Heart J.* **130**:1089-1093.

Yoshida M, Suzuki A, Yamamoto H, Noguchi S, Mizuno Y, Ozawa E. (1994) Dissociation of the complex of dystrophin and its associated proteins into several unique groups by n-octyl beta-D-glucoside. *Eur J Biochem*. **222**:1055-61.

Yount RG, Lawson D, Rayment I. (1995) Is myosin a "back door" enzyme? *Biophys J*. **68 (Suppl)**:44S-49S.

Yue TL, Wang C, Gu JL, Ma XL, Kumar S, Lee JC, Feuerstein GZ, Thomas H, Maleeff B, Ohlstein EH. (2000) Inhibition of extracellular signal-regulated kinase enhances Ischemia/Reoxygenation-induced apoptosis in cultured cardiac myocytes and exaggerates reperfusion injury in isolated perfused heart. *Circ Res.* **86**:692-699.

Zhu W, Shiojima I, Hiroi Y, Zou Y, Akazawa H, Mizukami M, *et al*. (2000) Functional analyses of three *Csx/Nkx2.5* mutations that cause human congenital heart disease. *J Biol Chem* **275**:35291-35296.

Zolk O, Flesch M, Nickenig G, Schnabel P, Bohm M. (1998) Alteration of intracellular Ca2(+)-handling and receptor regulation in hypertensive cardiac hypertrophy: insights from Ren2-transgenic rats. *Cardiovasc Res*. **39**:242-256.